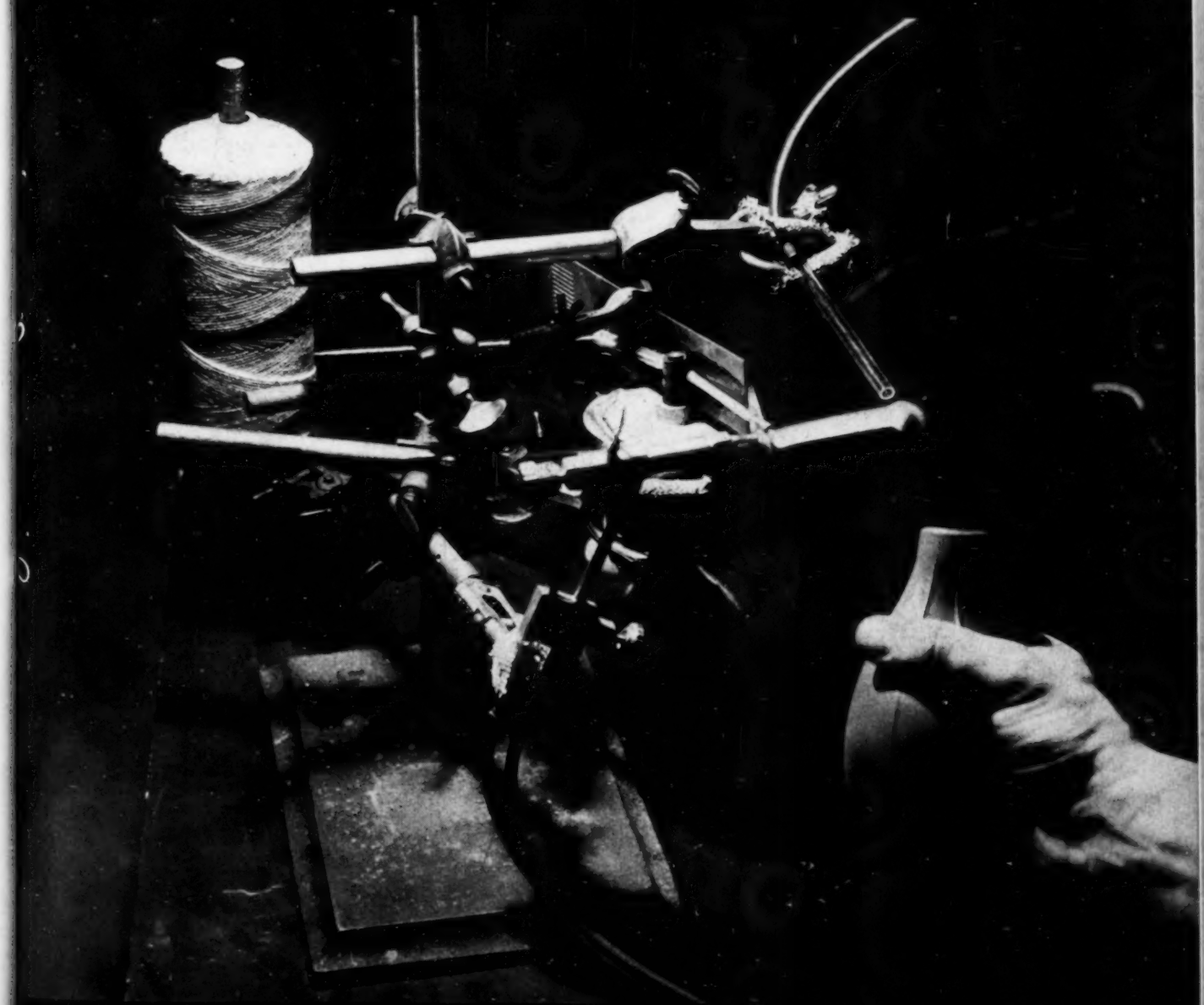
JOURNAL OF APPLIED PHYSICS



JOURNAL OF APPLIED PHYSICS

Vol. 18, No. 7 JULY, 1947

In This Issue

Page

Nucleation of Ice Formation by Silver Iodide B. Vonnegut 593

Waves in Elastic Tubes: Velocity of the Pulse Wave in Large Arteries Allen L. King 595

Electronic Computer for X-Ray Crystal Structure Analyses R. Pepinsky 601

Resonant Cavities for Dielectric Measurements C. N. Works 605

Variable Capacitor for Measurements of Pressure and Mechanical Displacements. A Theoretical Analysis and Its Experimental Evaluation John C. Lilly, Victor Legallais, and Ruth Cherry 613

Conductivity of Metallic Surfaces at Microwave Frequencies E. Maxwell 629

Note on Circular Loop Antennas with Non-Uniform Current Distribution G. Glinski 638

High Speed Compression Tests on Copper M. Greenfield and E. T. Habib 645

$TM_{0,1}$ Mode in Circular Wave Guides with Two Coaxial Dielectrics Sidney Frankel 650

Graphical Methods for Evaluating Fourier Integrals W. J. Cunningham 656

Two Calculating Machines for X-Ray Crystal Structure Analysis A. D. Booth 664

Field of a Microwave Dipole Antenna in the Vicinity of the Horizon C. L. Pekeris 667

Letters to the Editor:

Linearized Supersonic Flow through Ducts H. F. Ludloff and F. Reiche 681

Rigorous Solution of Linearized Supersonic Flow through Ducts C. E. Mack, Jr., H. F. Ludloff, and F. Reiche 682

Mechanism of Cutting Fluid Action Milton C. Shaw 683

In Every Issue

New Books 684

New Booklets 688

Here and There 689

Calendar of Meetings 690

Index to Advertisers xxx

BOARD OF EDITORS

Elmer Hutchisson, *Editor*

Associate Editors

R. Bowling Barnes
Otto Beeck
W. F. Busse
Selig Hecht
W. James Lyons
Philip Morrison
Paul E. Sabine
Frederick Seitz
J. R. Townsend
John Von Neumann

Publication of the American Institute of Physics

ADMINISTRATIVE STAFF

Henry A. Barton, *Director*

Cleveland Norcross, *Assistant Director*

Ruth F. Bryans, *Publication Manager*

Theodore Vorburger, *Business Manager*

John T. Tate, *Adviser on Publications*

GOVERNING BOARD

George R. Harrison, *Chairman*

Wallace Waterfall, *Secretary*

George B. Pegram, *Treasurer*

E. U. Condon
A. J. Dempster
W. F. Fair, Jr.
Floyd A. Firestone
Harvey Fletcher
R. C. Gibbs
K. S. Gibson
Paul E. Klopsteg
A. H. Pfund
John T. Tate
M. W. White
A. G. Worthing

The Journal of Applied Physics, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as Physics. The Journal of Applied Physics publishes editorials, news of physicists, and reviews, as well as technical papers of applied physics.

Manuscripts should be submitted to Elmer Hutchisson, Editor, Case School of Applied Science, Cleveland 6, Ohio.

Subscriptions and orders for back numbers should be addressed to Prince and Lemon Streets, Lancaster, Pa. or to the American Institute of Physics, 57 East 55 Street, New York 22, N. Y.

Subscription Price	U.S. and Canada	Elsewhere
To members and associate members of the American Institute of Physics	$6.00	$7.00
To all others	8.00	9.00

Back Number Prices

Complete set: Vol. 1, July, 1931–Vol. 17, 1946—$124.00

Yearly back number rate: $10.00

Single copies: $1.00 each.

Changes of address and proofs should be addressed to the Publications Manager.

Advertising rates supplied on request. Orders, advertising copy, and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879. Accepted for mailing at the special rate of postage provided for in the Act of February 28, 1925, authorized May 2, 1932.

Copyright 1947, by the American Institute of Physics.

Journal
of
Applied Physics

Volume 18, Number 7 July, 1947

The Nucleation of Ice Formation by Silver Iodide

B. VONNEGUT
General Electric Research Laboratory, Schenectady, New York
(Received March 17, 1947).

Silver iodide particles have been found to serve as nuclei for the formation of ice crystals in super-cooled water and in water vapor super-saturated with respect to ice. It is believed that silver iodide serves as a very effective nucleus because it very closely resembles ice in crystal structure. Both dimensions of the unit cell of ice and silver iodide are the same to within approximately one percent. The maximum temperature at which the silver iodide particles serve as nuclei is approximately −4°C for particles one micron in diameter, and −8°C for particles 100 Angstrom units in diameter. A silver iodide smoke generator has been constructed which consumes 1 mg of silver iodide per second and produces 10^{13} effective nuclei per second.

V. J. SCHAEFER of this laboratory has recently reported that super-cooled liquid water clouds spontaneously transform from water to ice when the temperature is −35°C or lower. He observed that at temperatures as low as −20°C, the addition of fine particles of a wide variety of substances was without effect on the super-cooled cloud.[1]

B. M. Cwilong,[2] using a Wilson cloud chamber, found that moisture in clean air is precipitated as a cloud of ice crystals if the temperature during the expansion drops to −35.0°C or lower and the air is super-saturated with respect to ice at this temperature. He observed that the ice crystals were formed even though the expansion ratio was far less than that necessary to produce the formation of liquid water drops. Cwilong observed that if "ordinary atmospheric air" is used, ice crystals form at −27°C, and that if the air is "artificially contaminated with tobacco smoke," the limiting temperature is −23°C. He found "crystalline dust" to be no more active than "amorphous dust" in producing ice crystals.

Following Schaefer's work, experiments were made using Schaefer's apparatus and technique to see whether particles of substances very similar to ice in crystal structure might not serve as nuclei for ice formation at temperatures closer to the freezing point.

A search was made through x-ray crystallographic data for substances resembling ice as closely as possible in crystal system, space group,

TABLE I.

Substance	System	Space group	Lattice constant	
Ice	hex.	D_{6h}^4	4.535	7.41*
AgI	hex. ZnO	C_{6v}^4	4.585_6	7.490**
PbI_2	hex.	D_{3d}^3	4.54	8.86

* William H. Barnes, Proc. Roy. Soc. London **A125**, 670–693 (1929).
** N. H. Kolkmeijer, W. J. D. von Dobbenburgh, and H. A. Boekehoogen, Proc. Amsterdam **31**, 1014–1027 (1928).

[1] V. J. Schaefer, "The production of ice crystals in a cloud of supercooled water droplets," Science **104**, 457–459 (1946).
[2] B. M. Cwilong, Nature **155**, 361–362 (1945).

and dimensions of unit cell. Two substances were chosen which are listed in Table I along with ice for comparison.

These crystalline substances were tested to see if they would act as nuclei for the formation of ice when dusted as a powder into a super-cooled water cloud in Schaefer's apparatus at a temperature of −20°C.

Silver iodide was without apparent effect; however, the introduction of lead iodide powder caused the formation of a number of ice crystals in the super-cooled cloud. It appeared that far fewer ice crystals formed than would have been expected from the number of particles of lead iodide introduced.

The supposition that the lead iodide particles acted as nuclei because one of the dimensions of its unit cell was almost the same as that for ice was discarded when V. J. Schaefer discovered that iodiform crystals and iodine vapor behave in a similar way. The orthorhombic structure of iodine bears little relation to the structure of ice.

Later, investigations were made on the nucleating effect of smokes produced by electric sparks between electrodes of various metals. It was discovered that a single spark between silver electrodes in the presence of iodine vapor produced many thousands of times as many ice nuclei as lead iodide particles or iodine vapor alone. It was then found that silver iodide smoke, produced by heating silver iodide on a hot filament or by dispersing it in a flame, produced enormous numbers of nuclei. The failure of the initial experiment using silver iodide powder is attributed to the fact that the sample used was badly contaminated with soluble salts, a fact not known at the time. Powdered silver iodide without this impurity has since been found to serve as nuclei for ice formation.

It is believed that silver iodide acts as a very effective nucleus for the formation of ice crystals because it very closely resembles ice in crystal structure. It can be seen that both dimensions of the unit cell of ice and silver iodide are the same to within about one percent.

According to Dr. D. Harker of this laboratory, the arrangements of the atoms in the unit cells of ice and of silver iodide are almost identical despite their different space groups. The structure of ice is the same as that of silver iodide with the oxygen atoms occupying positions corresponding to the silver and iodine atoms. In silver iodide, each silver atom is bonded tetrahedrally to four iodine atoms. In ice each oxygen atom is bonded tetrahedrally, through hydrogen bridges, to four oxygen atoms.

Iodine vapor alone, or fine silver particles alone are without appreciable effect in forming ice crystals at a temperature of −20°C. If the experiment is carried out using iodine vapor alone in a system closed to the atmosphere of the laboratory, iodine vapor soon loses its property of nucleating a super-cooled cloud. It is believed that iodine compounds and iodine vapor act to produce nuclei by reacting to form silver iodide with minute traces of silver in the laboratory atmosphere. Such traces of silver might be caused by sparks from electrical equipment using contacts made of silver or copper that contains silver. Exceedingly small amounts of silver introduced into a super-cooled cloud in the presence of iodine vapor cause the formation of very large numbers of nuclei. One breath of air blown over a silver wire heated red will produce many millions of ice nuclei if it is introduced into a super-cooled cloud containing a small amount of iodine vapor.

Experiments are being made on methods for the production of silver iodide smokes consisting of very large numbers of very small particles. The most effective method found thus far for the production of such smokes is the following. A cotton string coated with silver iodide is fed at a fixed rate into an oxyhydrogen flame which vaporizes the silver iodide. A few inches away from the point at which the silver iodide is introduced, the flame is rapidly quenched by blowing a strong jet of compressed air through it. The blast of air quickly cools and dilutes the vaporized silver iodide in the flame so rapidly that it condenses to form an invisible smoke of very small silver iodide particles. Measurements have been made which show that a generator of this sort produces ice nuclei at the rate of approximately 10^{13} per second with the consumption of one milligram of silver iodide per second. Mr. E. F. Fullam has examined particles of this smoke with the electron microscope and found

that they are of the order of 100 Angstrom units in diameter. One milligram, it can be computed, should yield about 3×10^{14} particles. The discrepancy between the two values may be caused by the fact that a portion of the silver iodide introduced into the flame may form a few very large particles or may form particles which for some reason are not effective nuclei.

It has been found that the effectiveness of the silver iodide particles as ice nuclei depends to a certain extent on the size of the particles. The smoke made by the oxyhydrogen flame technique has a particle size of the order of 100A diameter. This smoke does not cause large numbers of crystals to form until the temperature is -8°C or lower. Smokes made by vaporizing silver iodide from a hot wire have a particle size of the order of one micron and are effective at -4°C or below. A piece of silver iodide several millimeters in diameter placed in water in a test tube makes it difficult to super-cool the water to temperatures below about -3.5°C.

Recently the effect of silver iodide smoke was tried near Schenectady on a thin layer of supercooled fog at a temperature of -4°C. The smoke was produced by evaporating silver iodide on the surface of an electrically heated wire coil. The liquid water fog was transformed into small ice crystals for a distance of at least 150 feet from the point where the smoke was generated.

Waves in Elastic Tubes: Velocity of the Pulse Wave in Large Arteries

ALLEN L. KING
Dartmouth College, Hanover, New Hampshire
(Received February 28, 1947)

An equation for waves in elastic tubes is developed and applied to cylindrical tubes with Hookian and with elastomeric walls. The former application yields the Moens-Korteweg formula, which has been found inadequate. The latter application leads to a rather complicated equation for the pulse-wave velocity.

Values of pulse-wave velocities are computed for the thoracic aorta by means of the foregoing equation, and the results are compared with measured mean velocities for the entire aorta. Graphs and tables are given, so that a graphical analysis of this velocity equation can be applied to any large artery; and the method is illustrated by computing the pulse-wave velocity in the left common carotid. The relation $l \div \int_0^l ds/v$, for the mean velocity over a tube of length l, is shown to be valid for the aorta and large arteries.

IN a series of papers Hamilton, Remington, and Dow have published an extensive and critical survey of previous work and of their own important contributions to the study of propagation velocities for the arterial pulse wave[1] and its relation to stroke volume[2] and the cardiac ejection curve.[3] They found their results to be in accord with those of previous investigators. At best, only qualitative agreement was found between measured velocities and those computed from the Moens-Korteweg formula[4]

$$v = F(Ee_0/2\rho r')^{\frac{1}{2}}. \qquad (1)$$

In this relation E is Young's modulus of the wall material, e_0 is the thickness of the undistended wall, r' is the internal equilibrium radius of the tube, and ρ is the density of the fluid within the tube. Moens' factor F is an empirical dimension-

[1] W. F. Hamilton, J. W. Remington, and P. Dow, Am. J. Physiol. **144**, 521 (1945).
[2] J. W. Remington, W. F. Hamilton, and P. Dow, Am. J. Physiol. **144**, 536 (1945).
[3] J. W. Remington, and W. F. Hamilton, Am. J. Physiol. **144**, 546 (1945).

[4] Moens, A. I. *Die pulsekurve* (E. J. Brill, Leiden, 1878) p. 90; D. S. Korteweg Ann. d. Physik u. Chem. **241**, 525 (1878). Dr. H. J. Ralston of the College of Physicians and Surgeons of San Francisco has called my attention to a paper by Thomas Young [Trans. Roy. Soc. London **98**, 164 (1808)], wherein under the title, "Of the propagation of an impulse through an elastic tube," a formula for pulse-wave velocity is given, which can be transformed into that commonly ascribed to Korteweg.

less constant to which Moens assigned a value of 0.9, as a result of experiments with standing waves in elastic tubes. Hamilton and co-workers found the same value of Moens' factor for Gooch tubing; but for aortas they obtained better agreement with a value of 0.6–0.7 for the dog and 0.8 for the human, at diastolic pressure levels above 70 mm Hg. At lower pressure levels Moens' factor is not constant, and there is not even qualitative agreement.

In this paper a general wave equation is derived for the case of a homogeneous, non-viscous, and incompressible fluid in a thin-walled elastic tube. The Moens-Korteweg formula, with F equal to unity, is readily obtained by introducing a simple version of the pressure-radius relation for a Hookian elastic tube with walls of negligible thickness. Aortas, however, have elastic properties similar to those of rubber tubes and, therefore, the pressure-radius relation for tubes with elastomeric walls should be used.[5]

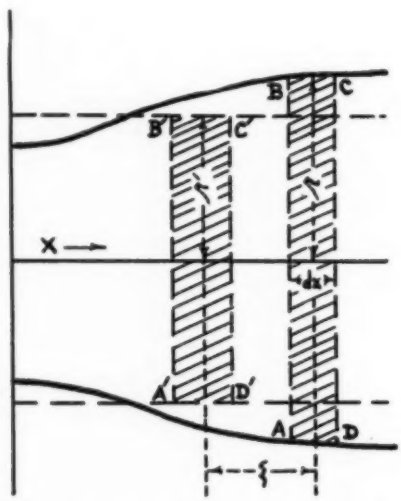


FIG. 1. Diagram of a portion of an elastic tube over which a wave is traveling.

Pulse-wave velocities computed from the new equation are in good agreement with measured values. Even so, it should be realized that arterial pulsations are larger and vessel walls are thicker than assumed here. And, furthermore, blood is neither homogeneous nor non-viscous.

[5] A. L. King, J. App. Phys. **17**, 501 (1946).

THE WAVE EQUATION

A homogeneous non-viscous incompressible fluid flows through a cylindrical tube with an elastic wall. The wall is sufficiently thin so that its inertia can be neglected in comparison with that of the fluid. The speed of a transverse wave down the tube is desired.

Let a thin element of fluid $ABCD$ of length dx and radius r be displaced a distance ξ from its undisturbed position $A'B'C'D'$ when a wave passes along the tube (Fig. 1). With no wave in the tube the radius is r', and the internal pressure is p'. At all times the external pressure is p_0.

The pressure on the fluid arises from the distension of the tube wall and is a function of the radius. Let the instantaneous excess pressure P at AB equal $(p-p')$, where p is the instantaneous absolute pressure at AB. Then the excess force on surface AB is $\pi r^2 P$, and the excess force on surface CD is $\pi r^2 P+(\partial/\partial x)(\pi r^2 P)\cdot dx$. The force accelerating the element $ABCD$ in the positive x-direction, therefore, is $-(\partial/\partial x)(\pi r^2 P)dx$. But, the mass of this fluid element is $\pi r^2 \rho dx$; so that the equation of motion can be written

$$-\frac{\partial}{\partial x}(r^2 P)=\rho r^2\frac{\partial^2\xi}{\partial t^2}. \tag{2}$$

The thickness of the element at $A'B'C'D'$ is $(1-\partial\xi/\partial x)dx$ and, since the fluid is incompressible,

$$r^2=r'^2(1-\partial\xi/\partial x). \tag{3}$$

Equation (2) can now be written

$$\frac{\partial^2\xi}{\partial t^2}=-\frac{1}{\rho}\left\{P(1+\partial\xi/\partial x)\frac{\partial^2\xi}{\partial x^2}-\frac{\partial P}{\partial x}\right\}. \tag{4}$$

This is the equation for the pulse-wave.

TUBES WITH HOOKIAN WALLS

For tubes with walls having a constant value of elastic modulus and following Hooke's law, Korteweg used the simple relation in which the wall thickness is assumed not to change. For this case the instantaneous excess pressure

$$\begin{aligned}P&=Ee_0(1/r'-1/r)\\&=-(Ee_0/2r')(\partial\xi/\partial x).\end{aligned} \tag{5}$$

On inserting this value of P in the general equa-

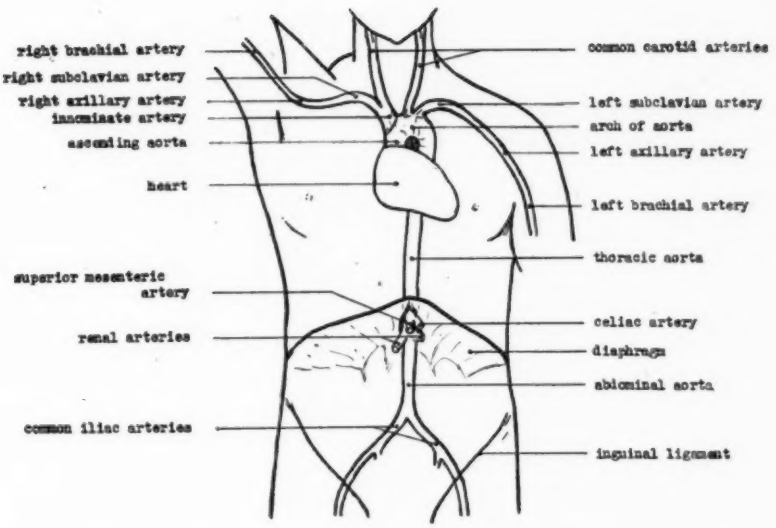


FIG. 2. Schematic drawing showing the approximate positions of the aorta and some of its branches.

tion for the pulse-wave and on neglecting terms involving the square of $\partial\xi/\partial x$, the wave equation,

$$\frac{\partial^2\xi}{\partial t^2}=(Ee_0/2\rho r')\frac{\partial^2\xi}{\partial x^2}, \tag{6}$$

is obtained. The velocity of the wave is

$$v=(Ee_0/2\rho r')^{\frac{1}{2}}, \tag{7}$$

namely, the Moens-Korteweg formula with F equal to unity.

In general, the thickness of the tube wall does not remain constant. If this effect is taken into account, the instantaneous excess pressure is given by

$$\begin{aligned}P&=2e_0E(r-r')/(2-\sigma)r_0^2\\&=-\{Ee_0r'/(2-\sigma)r_0^2\}(\partial\xi/\partial x),\end{aligned} \tag{8}$$

where σ is Poisson's ratio of the wall material. When this value of P is used in Eq. (4) and terms involving the square of $\partial\xi/\partial x$ are neglected, the wave velocity comes out to be

$$v=\{Ee_0r'/\rho r_0^2(2-\sigma)\}^{\frac{1}{2}}. \tag{9}$$

This relation for pulse-wave velocity yields a somewhat higher value than that obtained from Eq. (7), since r' is greater than r_0 and σ ordinarily is positive.

TABLE I. Values of A and β for several age groups.

Age range (years)	20–24	29–31	36–42	47–52	71–78
A (mm Hg)	88	74	70	65	44.5
β	0.302	0.415	0.462	0.510	0.640

TUBES WITH ELASTOMERIC WALLS

Elastomers are rubber-like substances containing long chains of molecules. They are highly resilient and have the unusual thermoelastic property of contracting on heating, when they are under tension. Living tissues have these characteristics and, in particular, the walls of blood vessels may be considered elastomeric.

For a thin-walled elastomeric tube the instantaneous excess pressure is

$$\begin{aligned}P=A[&(r_0/r)^{\frac{1}{2}}\{\mathfrak{L}^{-1}(\beta r/r_0)/\mathfrak{L}^{-1}(\beta)\}\\&-(r_0/r')^{\frac{1}{2}}\{\mathfrak{L}^{-1}(\beta r'/r_0)/\mathfrak{L}^{-1}(\beta)\}\\&+(r_0/r')^{\frac{1}{2}}-(r_0/r)^{\frac{1}{2}}],\end{aligned} \tag{10}$$

where A represents $e_0p_0/2r_0$, and β is equal to the ratio of the circumference of the tube to the maximum length of a molecular chain.[5] For arteries β is an age-dependent parameter. The functions $\mathfrak{L}^{-1}(\beta)$, $\mathfrak{L}^{-1}(\beta r/r_0)$, and $\mathfrak{L}^{-1}(\beta r'/r_0)$ are inverse Langevin functions.

For small variations in the radius, as the wave progresses along the tube, $\mathfrak{L}^{-1}(\beta r/r_0)$ may be expanded about r', and only the first two terms need be retained. The result is

$$\mathfrak{L}^{-1}(\beta r/r_0)=u'\{1-\tfrac{1}{2}(Br'/r_0)(\partial\xi/\partial x)\}, \tag{11}$$

where

$$B=\beta u' \sinh^2 u'/(\sinh^2 u'-u'^2) \quad (12)$$

and

$$u'=\mathfrak{L}^{-1}(\beta r'/r_0). \quad (13)$$

Eq. (10) now reduces to

$$P=\tfrac{1}{4}A(r_0/r')^{\frac{1}{2}}\{(1-2Br'/r_0)(u'/u_0) - 3(r_0/r')^{\frac{1}{2}}(\partial\xi/\partial x), \quad (14)$$

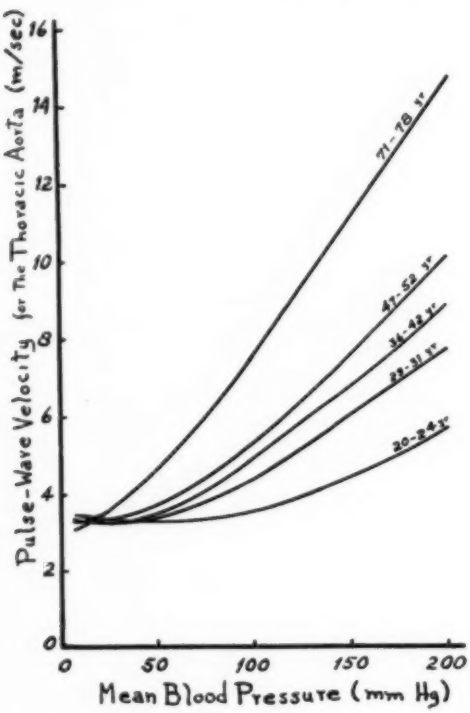


FIG. 3. Computed values of pulse-wave velocity for the human thoracic aorta, showing the dependence on age and mean blood pressure.

in which $u_0=\mathfrak{L}^{-1}(\beta)$. The ratio u'/u_0 may be eliminated by use of the relation (Eq. (10) in reference 5).

$$p'-p_0=A\{(r_0/r')^{\frac{1}{2}}(u'/u_0)-(r_0/r')^{\frac{1}{2}}\}. \quad (15)$$

Insert the resultant expression for P in Eq. (4), neglect terms involving the square of $\partial\xi/\partial x$, and solve for the velocity of propagation of a pulse-wave in the elastomeric tube. The velocity is given by the equation

$$v^2=\tfrac{1}{4}(A/\rho)[(2Br'/r_0-1) \times\{(p'-p_0)/A+(r_0/r')^{\frac{1}{2}}\}+3(r_0/r')^{\frac{1}{2}}]. \quad (16)$$

PULSE-WAVE VELOCITIES IN THE HUMAN AORTA

Equation (16) for pulse-wave velocity may be applied to blood vessels. Note that the velocity increases with a rise in mean pressure within the tube. Also it increases with age, for the parameter β in the expression for B is age-dependent. Furthermore, it is higher when the ratio of wall thickness to internal diameter is larger, primarily through the factor A. All these conclusions are in qualitative agreement with observations on arteries of dog and man.

For a quantitative study, pulse-wave velocities are computed for that part of a human thoracic aorta just beyond the arch (Fig. 2). Values of A and β for several age groups were evaluated elsewhere[5] and are given in Table I. The specific gravity of blood is assumed to equal 1.057 for all ages.[6] The resulting computed velocities at pressures of 7.5, 50, 100, 150 and 200 mm Hg are gathered in Table II and are plotted in Fig. 3. Near zero pressure, the velocity is seen to decrease as the aorta ages. At the normal ranges of aortic pressures, however, there is a general increase of velocity with age. This increase becomes much pronounced in the hypertensive range. As noted in Fig. 3, the pulse-wave velocity is very nearly a linear function of pressure above 120 mm Hg.

Ordinarily, experimental determinations of pulse-wave velocities are made over long arterial paths. For instance, Hallock[7] made an extensive statistical study of aortic velocities in nearly 500 humans ranging in age from 5 to 85 years. His values of mean velocities are based on measurements of differences in distances and propagation times for pulse-waves traveling over the arch, thoracic and abdominal aortas, and the left iliac artery to the inguinal ligament and over the left common carotid artery (Fig. 2). To compute such mean velocities it would be necessary to know the velocity over every element of length along

TABLE II. Pulse-wave velocities for the thoracic aorta in meters per second.

Age (years)	Pressure (mm Hg) 7.5	50	100	150	200
20–24	3.47	3.29	3.57	4.50	5.74
29–31	3.31	3.36	4.38	6.11	7.88
36–42	3.28	3.49	4.91	6.70	8.94
47–52	3.26	3.71	5.32	7.53	10.16
71–78	3.02	4.65	7.57	11.12	14.73

[6] W. F. Hamilton, Glasser's *Medical Physics* (Year Book Publishers, Chicago, 1944) p. 115.
[7] P. Hallock, Arch. Internal Medicine **54**, 770 (1934).

the path. In such cases, the mean velocity $\bar{v}$ is given by the relation

$$\frac{1}{\bar{v}}=\frac{1}{l}\int_0^l \frac{ds}{v}, \qquad (17)$$

where l is the path length. If aging is uniform over the entire aorta, the ratio of mean velocity to the velocity in the thoracic aorta alone should be independent of age. Furthermore, the ratio should be greater than unity, since the pulse-wave velocity increases as the vessels become smaller in diameter and their walls thicken. On comparing the theoretical values of velocities for the thoracic aorta at a mean blood pressure of 100 mm Hg with Hallock's measured mean velocities, a constant empirical factor of 1.37 can be found. The approximate agreement between the adjusted theoretical values and Hallock's data is shown in Fig. 4.

PULSE-WAVE VELOCITIES IN OTHER LARGE ARTERIES

Bazett, Cotton, Laplace, and Scott[8] published measured values of average pulse-wave velocities for paths from the heart, through the ascending aorta, arch, and innominate artery, to the subclavian artery, and also from the subclavian, through the axillary, to the branchial artery (Fig. 2). For the three men in the age-range of 20–24 years, whom they investigated, the mean

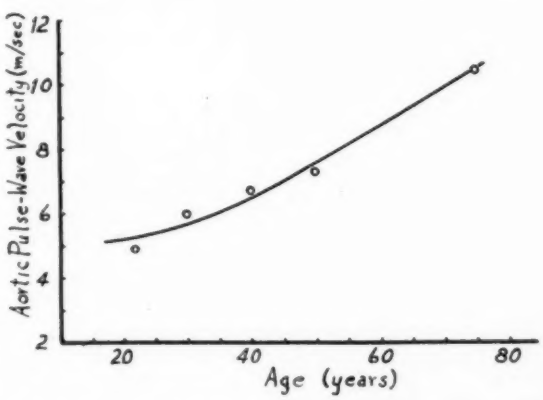


FIG. 4. The curve shows Hallock's values of mean pulse-wave velocities for the entire aorta as a function of age. The points were obtained by multiplying the theoretical values of velocity for the thoracic aorta alone, at a mean blood pressure of 100 mm Hg, by the empiric factor 1.37.

[8] H. C. Bazett, F. S. Cotton, L. B. Laplace, and J. C. Scott, Am. J. Physiol. **113**, 312 (1935).

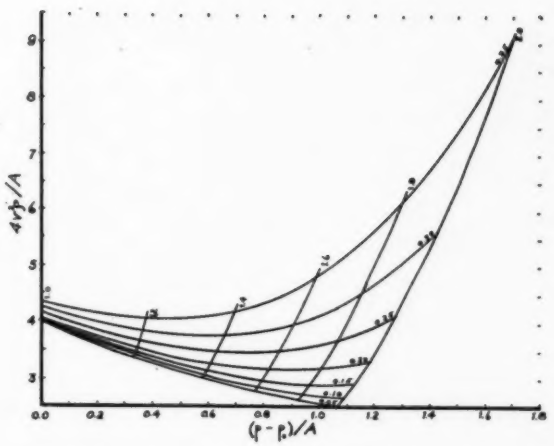


FIG. 5. Graph of $4v^2\rho/A$ as a function of $(p-p_0)/A$ for several values of β and r/r_0. Values of β range from 0.05 to 0.35 in steps of 0.05, and values of r/r_0 range from 1.0 to 2.0 in steps of 0.2.

values of these velocities were 3.6 and 5.4 meters/sec., respectively, with less than 10 percent deviation. To obtain the mean velocity between heart and brachial artery from these data, Eq. (17) may be modified. Since the distances from heart to subclavian artery and from subclavian to brachial artery, as described by Bazett and co-workers, are approximately equal, the reciprocal of the mean velocity between heart and brachial artery will be equal to half the sum of the reciprocals of the two measured velocities. This computation yields the value 4.3 meters/sec.

The foregoing result may be compared with the observations of Bazett and Dreyer.[9] These investigators measured the pulse-wave velocity between heart and brachial artery of eight medical students with ages in the range 20–24 years. The mean velocity was equal to 4.1 meters/sec. at a mean blood pressure of 88 mm Hg with an average deviation of ±0.4 meter/sec.

Since the average dimensions of the arteries from aorta to brachial artery are equivalent to those of the common carotid, it may well be expected that the pulse-wave in the carotid should have a velocity of approximately 4 meters/sec. An extension of the analysis in the early part of this paper can be used to compute the velocity in any large artery. It will be applied to the common carotid artery.

[9] H. C. Bazett and N. B. Dreyer, Am. J. Physiol. **63**, 94 (1922).

GRAPHICAL ANALYSIS OF THE VELOCITY EQUATION

In Eqs. (15) and (16) r'/r_0 and β may be considered parameters for the quantities $(p'-p_0)/A$ and $4\rho v^2/A$. There is shown in Fig. 5 a family of curves for which β varies between 0.05 and 0.35 in steps of 0.05, and r'/r_0 varies between 1.0 and 2.0 in steps of 0.2. Other curves may be drawn for values of β and r'/r_0 outside these ranges, but for the following computations the curves of Fig. 5 suffice.

In order to use these curves, first A is evaluated from known values of external pressure and ratio of effective wall thickness to internal diameter of the undistended artery. Then from the mean blood pressure within the artery $(p'-p_0)/A$ is computed. A value for β next is found by making the reasonable assumption that, for the walls of the large arteries within a given body, the molecular chains all have the same length. For all such arterial walls, therefore, β is directly proportional to the internal diameter of the undistended vessel. Since values of β already are known for the upper part of the thoracic aorta (Table I), β for any other large artery may be found by the relation

$$\beta\,(\text{artery}) = \beta\,(\text{thoracic aorta}) \cdot \frac{d_0(\text{artery})}{d_0(\text{thoracic aorta})}. \quad (18)$$

From these values of β and $(p'-p_0)/A$ a value of $4\rho v^2/A$ may be read off the curves of Fig. 5 by interpolation and the pulse-wave velocity v can then be found.

To illustrate the procedure, a value for the pulse-wave velocity in the common carotid artery of a 20–24 year old person will be computed. From Kani's data[10] the mean ratio of wall thickness to internal diameter is estimated to equal 0.22 for twenty-year olds. For an external pressure p_0 of 770 mm Hg, A nearly equals 170 mm Hg. The average pressure within an artery near heart level is assumed to be the mean of the systolic and diastolic pressures. A fifteen percent error in this mean introduces less than two percent error in the computed velocity of the pulse-wave for the carotid artery. As noted earlier, the mean blood pressure for a 20-year old person nearly equals 90 mm Hg. The function $(p'-p_0)/A$, therefore, is equal to 90/170 or 0.53.

The ratio d_0(carotid)/d_0(thoracic aorta) for 20-year olds approximately equals one-third, so that for the carotid β equals 0.10. From Fig. 5, the quantity $4\rho v^2/A$ is found to have the value 3.15 at values of β and $(p'-p_0)/A$ equal to 0.10 and 0.53, respectively. But for whole blood the density ρ equals 1.057 g/cc, and here A equals 170 mm Hg, or 226,000 dynes/cm^2. From these data v is computed to be 4.1 meters/sec.

The foregoing method may be modified so as to obtain values of β from measurements of pulse-wave velocity and mean blood pressure.

No doubt other factors than those considered in this development may influence pulse-wave velocities in living systems. The experiments reported by Hamilton and co-workers[1] in their study of this problem, however, suggest that the propagation of a pulse wave along the aorta and large arteries depends almost completely on the mechanical properties of the intact living artery.

[10] I. Kani, Virchow's Arch. **201**, 45 (1910); see also A. L. King, Science **105**, 127 (1947) for comments on Kani's data.

An Electronic Computer for X-Ray Crystal Structure Analyses*

R. Pepinsky
Department of Physics, Alabama Polytechnic Institute, Auburn, Alabama
(Received January 29, 1947)

An electronic synthesizer is described for determination of atomic positions in crystals. The synthesizer sums the two-dimensional Fourier series representing planar, centro-symmetric projections of electron densities in a crystal unit cell; and the projection is presented by a television scan on the screen of a cathode ray oscilloscope. The specific advantage of the device is the immediate observability of effects on the projection of alterations in signs of one or any number of Fourier coefficients.

A METHOD has been developed for rapid visual presentation of the positions of atoms in a crystal structure by supplying x-ray diffraction data to an electronic computer coupled to a cathode ray oscillograph. The computer sums the two-dimensional Fourier series representing the projection on a lattice plane of the electron densities in a crystal unit cell; and the projection is presented by means of a television type of scan on the screen of the cathode ray tube. The method is based upon Huggin's modification[1] of Bragg's earlier photographic summation of simulated interference fringes[2]; in this respect it is related to the sand-deposition method of fringe summation reported by McLachlan and Champayne[3] in the December, 1946 issue of this journal. In contrast to these modifications of the Bragg method, however, accurate projections with various axial ratios and angles are readily permitted in the present computer.

Artificial interference fringes for an (*hk*0) projection are produced on the oscilloscope screen by applying to the CR tube grid intensity-modulating sinusoidal voltage signals of amplitude proportional to F_{hk0} and frequency f_{hk0}, the latter being given by the relation:

$$f_g = f_{hk0} = h \cdot f_V + k \cdot f_H, \qquad (1)$$

where f_V and f_H are the vertical and horizontal sweep frequencies respectively. By way of illustration: f_V may be a slow sweep, at 1 c.p.s., and f_H a fast sweep at 1000 c.p.s., and the duty cycle of both sweeps practically 100 percent. One 1000-line scanning frame then appears on the screen each second. The signals for various (*hk*0) fringes are mixed before application to the grid, so that a two-dimensional pattern occurs for each frame. The fluorescent screen is a long-persistence type (3- to 5-second decay period) to permit visualization, and the pattern may be photographed for permanent recording.

Only patterns with centers of symmetry are attempted. The *signs* of the structure factors F_{hk0}, which in general are not directly determinable from the x-ray data, are introduced by having each signal available in 0 or 180° phase, and selecting one or the other of these by means of a double-throw switch for each term. Amplitudes are adjusted by means of individual potentiometers for each factor.

It will be seen from Eq. (1) that when $h=0$, so that f_g is some integral multiple of the fast sweep, vertical fringes result on the screen. *Increasing* f_g slightly, with the sweep frequencies unchanged, causes the fringes to tilt (but remain parallel to one another) in one direction, because the intensity-modulating signal *advances* in phase with respect to the fast sweep. *Decreasing* the modulating frequency produces parallel fringes oppositely tilted to the first case. Certain refinements of this simple procedure, involving the use of lower duty-cycle sweeps but of the same sweep speed as indicated above, are required to maintain accurate placement of the fringes. These and other details of the technique will be described and published elsewhere.

A fast sweep of 1000 c.p.s. in the horizontal direction and a slow of 1 c.p.s. vertically have been utilized, and precise synchronization of

* Presented in part before the American Society for X-Ray and Electron Diffraction, Lake George, New York, June 11, 1946.

[1] M. L. Huggins, Jr., J. Am. Chem. Soc. **63**, 66 (1941).

[2] W. L. Bragg, Zeits. f. Krist. **70**, 475 (1929).

[3] D. McLachlan, Jr., and E. P. Champaygne, Jr., J. App. Phys. **17**, 1006 (1946).

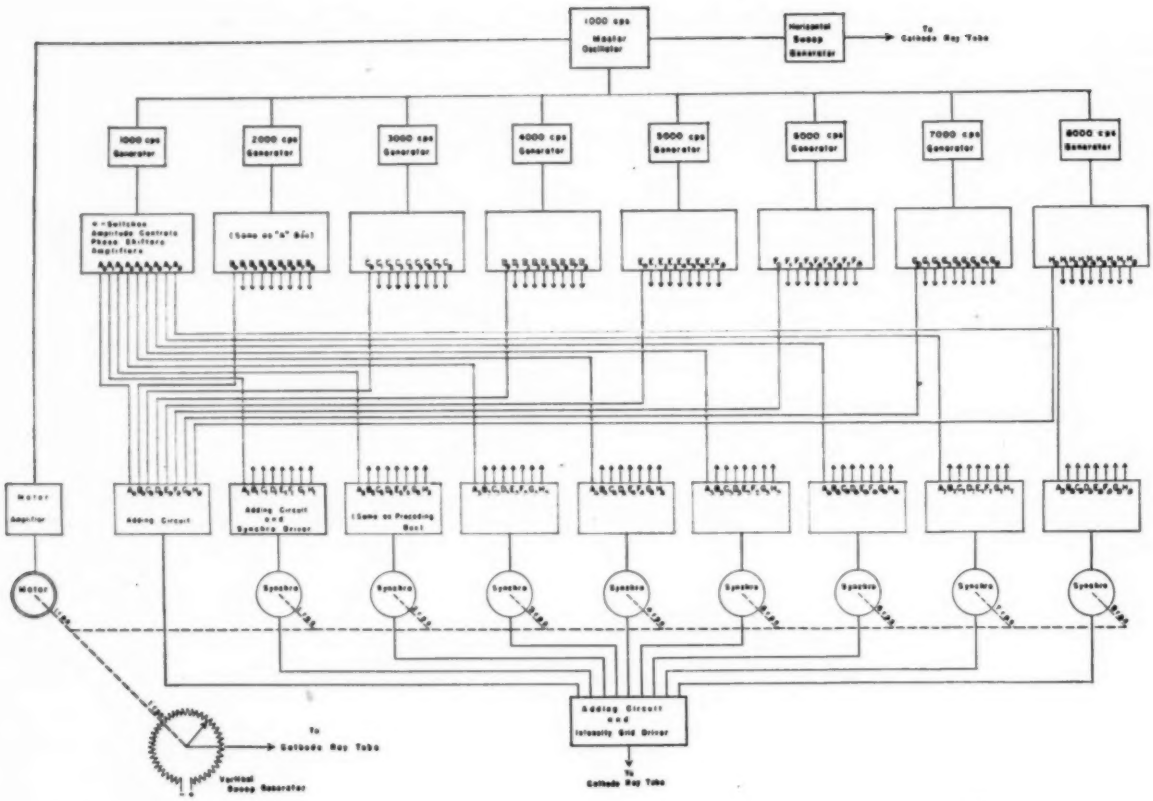


FIG. 1. Block diagram of Fourier synthesizer for (h, k) terms to (8, 8).

these with the grid signals, according to equation (1), is essential. All frequencies are derived from a 1000 c.p.s. base oscillator, which should be fairly frequency stable but need not be strictly so. The fast sweep is synchronized to this; the slow is obtainable through the use of a 1000 cycle motor with a 1 r.p.s. shaft, the motor being driven from a power amplifier.** A series of generators of frequencies $1000 \cdot k$ c.p.s. are synchronized by a heterodyning system to the base oscillator. Phase and amplitude-adjustment controls are available for each of these.

The key to the entire computer is the method for production of the frequencies $1000 \cdot k \pm h$ c.p.s. These must be accurately related to the base oscillator and to the fast and slow sweep speeds; no drifting whatever from the precise relation can be tolerated. The exactly-related frequencies are obtained as follows. Each $1000 \cdot k$ c.p.s. signal is fed to an integrating feed-back amplifier which produces a 90° phase-shifted signal of the original frequency. Application of these two signals—of equal amplitude and frequency but 90° phase-separation—to the stator windings of a two-phase selsyn produces a rotating electric field in the selsyn of rotational frequency $1000 \cdot k$ c.p.s. If the selsyn rotor remains stationary, a signal of frequency $1000 \cdot k$ c.p.s. appears on its leads. But if the rotor is rotated at h r.p.s. in a direction counter to the field, a signal of frequency $1000 \cdot k + h$ c.p.s. appears on the rotor leads. Rotation of the rotor at h r.p.s. in the *same* direction as the field rotation results in a signal of frequency $1000 \cdot k - h$ c.p.s. The rotor speeds are obtained by gearing in ratios 1:1, 1:2, $\cdots$1:h to the 1 c.p.s. shaft driven by the 1000-cycle motor.

All signals with the same h-value can be *mixed* and fed to a *single* selsyn, the rotor speed of which is h r.p.s. The number of selsyns required for the device is thus considerably less

** It will be demonstrated in a later paragraph that some of the synchronization here is redundant. Specifically, the motor need not be synchronized to the fast sweep signals and the $1000 \cdot k$ oscillators.

than the number of F_{hk0} terms. The 90° phasing amplifiers are made rather frequency-insensitive, so that only one such amplifier is required for each selsyn; this again reduces the number of components. Because the output signal from the phasing amplifiers will vary with *large* variations in signal frequency, normalizing potentiometers are supplied for each of the 1000·k c.p.s. oscillators.

A 1000·k+h c.p.s. signal is produced, as described, by rotation of the selsyn rotor in a direction counter to the rotating field direction. It is possible to obtain both a 1000·k+h c.p.s. and a 1000·k−h c.p.s. signal from a *single* selsyn by producing rotating fields in *opposite* directions in the same selsyn stator. To accomplish this it is necessary only to alter the phase of one 90° component of one or the other of the 1000·k c.p.s. signals by 180°. (In general, the amplitude of the 1000·k c.p.s. signal which will become a 1000·k+h c.p.s. frequency is different from the 1000·k c.p.s. signal which will become the 1000·k−h c.p.s. frequency, the amplitudes being controlled by the $(h, k, 0)$ and $(h, \bar{k}, 0)$ potentiometers, respectively.) Thus, for a matrix of (20, ±20) terms, only twenty selsyns are required.

A partial block diagram for a system of F_{hk0} terms up to (8, 8, 0) is shown in Fig. 1. The use of lower than 100 percent duty-cycle sweeps has not been indicated in this diagram, but such modifications do not alter the general operation of the system. The illustrated system achieves completely locked-in synchronization of all signals. Once the 1000·k c.p.s. signals have been phased so that their maxima occur at the start of the fast sweep, no drifting of fringes is possible.

A sample fringe, corresponding to F_{170}, is illustrated in Fig. 2. The fast sweep, at 1000 c.p.s., is from left to right in the figure; and the slow sweep, at 1 c.p.s., is from top to bottom. The grid signal here has the frequency 6999 c.p.s., as required by the formula:

$$f_{170} = 1 \cdot h + 1000 \cdot k = 1 \cdot -1 + 1000 \cdot 7 = 6999 \text{ c.p.s.}$$

The resulting pattern ordinarily will not show any discrete horizontal lines, because the distance between adjacent fast sweep traces is less than the breadth of the electron beam at its finest focus. In order to indicate the direction of the fast sweep here, an additional 60 c.p.s. modulation has been applied to the CR tube grid.

It can be shown, as suggested in the footnote on page 3, that there is a redundancy in the synchronization employed here. It is true that all the 1000·k oscillators and the fast sweep must be synchronized—and this is readily achieved. It is also necessary that the slow sweep and the selsyn rotations be synchronized, and this is achieved through the use of the potentiometer driven from the (approximately) 1 c.p.s. shaft and with the selsyns driven from the shafts geared to this. But it is *not* necessary that the shaft speeds, at multiples of about 1 r.p.s., be synchronized with the 1000 c.p.s. oscillator. If the shaft speed slows down, more horizontal sweeps occur per frame, and the phase-shift of each fringe signal is reduced; but these two effects exactly compensate one another, and the tilt of each fringe remains unaltered. This simplifies the circuitry, and permits the use of a standard (60-cycle) motor.

The fast and slow sweep directions are ordinarily at right angles. By adding to the fast sweep signal a *component* of the slow sweep signal, the angle between the effective slow sweep direction and the fast sweep can be varied over a wide range. The slow sweep component is obtained from another linear potentiometer on the same 1 c.p.s. shaft as the main slow sweep potentiometer. The amplitude of the slow com-

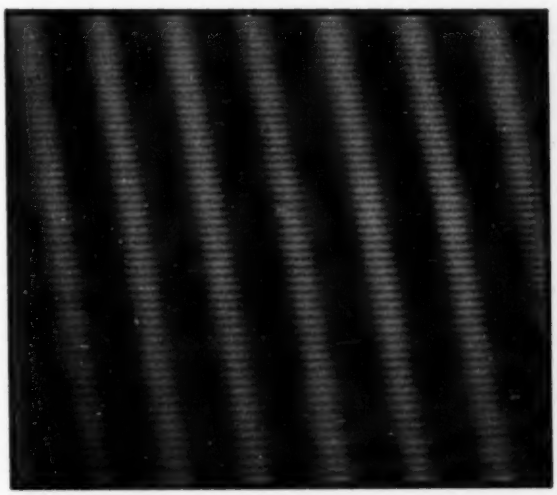


FIG. 2. $F_{\bar{1}70}$ fringe, with added 60 c.p.s. modulation to show fast sweep direction.

ponent is varied by means of the voltage applied across its potentiometer. Variation of the voltage across the main slow sweep potentiometer alters the axial ratio of the projection. Thus cell projections of any desired axial angle and ratio are attainable.

The computer can be built up in stages from one accommodating only a few F_{hk0} terms to one including as large a range of terms as appears necessary. It is intended that terms to (20, ±20, 0) be incorporated into the final synthesizer.

The basic difficulty in general crystal structural analyses arises from the absence of information concerning the phases of the structure factors. Whenever a center of symmetry occurs in a structure or in some projection of the structure (e.g., a projection perpendicular to a two-fold symmetry axis is centro-symmetric), the factors involved in an electron density map for the projection are known except for their algebraic sign. Some partial information as to these signs may be available from auxiliary information (atomic radii and bond angles as determined from related structures, Patterson projections showing interatomic vectors, optical data indicating group or molecular orientations, etc.); and in special cases the signs may be directly determinable.[4] In general, however, trial and error procedures are required. Since the order of from one hundred to several hundred factors are involved in a projection, the number of permutations of signs is overwhelming; and when tedious computations must be carried through each time a change is made, without a great deal of guidance in assignment of the signs structure determinations become unfeasible.

In order to reduce the work involved in the calculations, a good number of computational aids have been developed for structural analyses. These will not be reviewed here. It suffices to say that none of them meet in a satisfactory manner the peculiar requirements of this science, that alterations of signs of the factors be enterable in a simple way into the computer, and the effects of the alterations be very rapidly observable.

It is precisely this requirement of structural analyses which the present synthesizer is designed to satisfy. The signs of individual factors can be introduced and altered at will by simply throwing a switch for each factor; and the "summation" of a two-dimensional series is accomplished within a few seconds after the data is fed into the machine. The pattern corresponding to a specific choice of signs remains on the oscilloscope screen as long as is desired. The effects of changes in signs of one or any number of factors appear almost instantly, and the new pattern remains on the screen until another change is introduced. (The amplitudes of individual factors are equally adjustable, with the effects of such changes just as rapidly observable.) *Thus for the first time systematic variations of structure factor signs, in trial and error procedures, become feasible in a general structure analysis.*

Another advantage of the oscillographic presentation is the ease with which the *contrast* and *background level* of the pattern is alterable. The former is controllable through amplitude control of the final grid-modulating signal; and the background level is controlled through variation of the CR tube grid bias. Since the grid-modulating signals are mixed before their application to the cathode-ray tube, a non-linear relation between the beam current and the light-output of the fluorescent screen does not lead to errors in positions of peaks in the electron-density map. The size of a projection is also controllable, and any section of a projection can be selected and enlarged at will.

The computer should not now be considered as a replacement for accurately calculated contour maps of electron densities in the final stages of a successful structural determination. It is intended to facilitate the deduction of general features of a structure and to avoid the tediousness of trial and error stages of the analysis. It is conceivable that with proper care a device of this sort will be capable of quantitative results, but such is not its prime purpose.

An earlier form of this instrument was described by the writer at the Lake George, New York, meeting of the American Society for X-Ray and Electron Diffraction, June 11, 1946, and oscilloscope fringe patterns were shown there. An extended discussion of the system will be published elsewhere.

[4] J. M. Robertson, Reports on Progress in Physics **4**, 332 (1937).

Resonant Cavities for Dielectric Measurements*

C. N. Works
Research Laboratories, Westinghouse Electric Corporation, East Pittsburgh, Pennsylvania
(Received March 3, 1947)

Fixed and variable length re-entrant resonant cavities designed for the measurement of dielectric constant and dissipation factor are described. These cavities operate in the frequency decade of 10^8 to 10^9 cycles per second, a region avoided by many experimenters because the frequency is too high for the application of circuit techniques and not high enough for the convenient use of coaxial lines or wave guides. The theoretical considerations in the design of these cavities are presented. The well-known susceptance variation method, widely used in the frequency range of 10^4 to 10^8 cycles per second, was extended to apply to these cavities; it yields a rapid measuring technique and very simple expressions for calculating the values of the dielectric properties. The performance of this apparatus is discussed and the results of the measurements of a few typical dielectrics are given.

I. INTRODUCTION

DIELECTRIC constants and losses of materials had been measured even at ultra-high frequencies as early as 1895.[1] The early techniques were relatively crude compared with those developed within the last few years. Coaxial transmission line and wave guide methods, discussed by von Hippel,[2] Lamont,[3] Englund,[4] and Benoit,[5] in which a section of (usually resonent) coaxial line or wave guide is filled with dielectric, are used at present, especially at frequencies greater than 1000 megacycles per second.

The frequency band of 100–1000 megacycles has been avoided by most experimenters in dielectric measurements because the frequency is too high for the application of the usual circuit techniques and not high enough for the convenient use of the coaxial line or wave guide. The search for a simple and rapid method of making measurements in this frequency range has resulted in the development of two types of re-entrant cylindrical cavities in which small disk-shaped samples are used. This paper will be concerned with the description of these cavities as well as with their theory and performance.

II. DESCRIPTION OF THE APPARATUS

1. The Fixed Cavity

This instrument, which is shown in Fig. 1, operates in the frequency range of 400–600 megacycles per second. It is similar to an early model developed for the frequency range of 100–300 megacycles,[6] except that as a result of

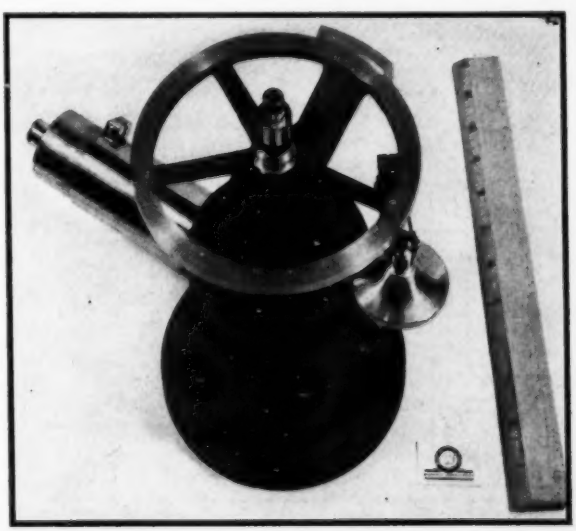


FIG. 1. The fixed cavity.

* The variable length cavity described in this paper was the subject of a thesis submitted to the Graduate School of the University of Pittsburgh, in partial fulfillment of the requirements of the M. S. degree.

[1] Drude, Wied Ann. **55**, 633 (1895); Zeits. f. physik. Chemie **23**, 267 (1897).

[2] S. Roberts and A. von Hippel, "New Method for Measuring Dielectric Constant and Loss in the Range of Centimeter Waves," J. App. Phys. **17**, 610–616 (1946).

[3] H. R. L. Lamont, "Theory of Resonance in Microwave Transmission Lines with Discontinuous Dielectric," Phil Mag. **29**, 521 (1940).

[4] C. R. Englund, "Dielectric Constants and Power Factors at Centimeter Wave Lengths," Bell Sys. Tech. J. **23**, 114–129 (1944).

[5] J. Benoit, J. de phys. et rad. **5**, 173–184, 203–216 (1944).

[6] C. N. Works, T. W. Dakin, and F. W. Boggs, "A Resonant Cavity Method for Measuring Dielectric Properties at Ultra High Frequencies," Proc. I.R.E. 245–254 (April 1945), or Trans. A.I.E.E. **63**, 1092–1098 (1944).

experience with this earlier instrument, a number of mechanical improvements have been made.

The sample required is a disk one inch in diameter and from 0.050 to 0.250 inch thick. It is inserted through the port near the bottom and placed between the re-entrant posts or electrodes for the measurement of its dielectric constant and dissipation factor.[7] This port is located in a portion of the cavity where the current is small. It is closed by a tapered plug. It has been found that the resonant frequency of the cavity is not altered by removal and replacement of the plug. The upper electrode is movable and is connected with the coaxial portion of the cavity by a sylphon bellows which eliminates all sliding contacts. The separation between these electrodes is controlled by a precision micrometer (Van Keuren Company) with a six-inch dial. The constructional details of this cavity are shown in Fig. 2.

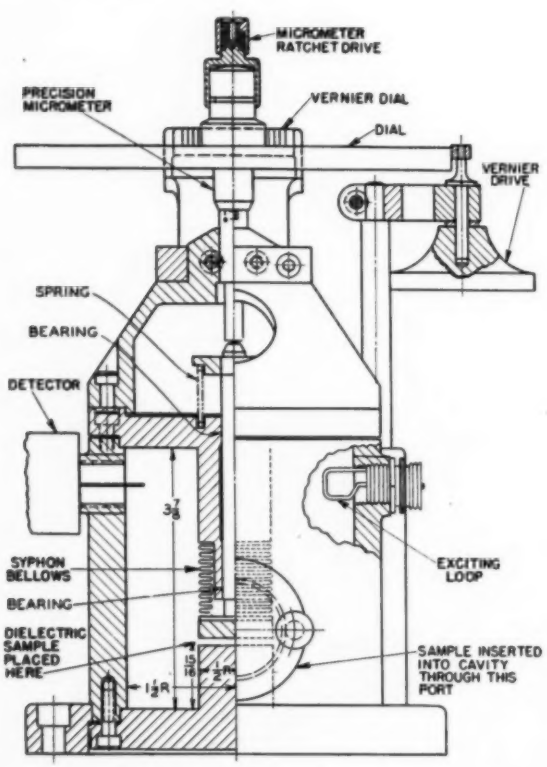
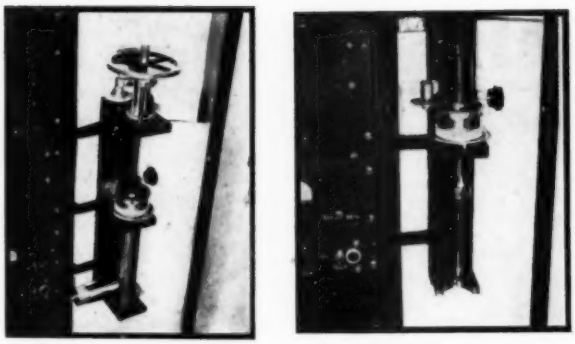


FIG. 2. 400–600-Mc dielectric measuring apparatus.

[7] Dissipation factor ($\tan\delta$) is the tangent of the loss angle. This angle is the complement of the phase angle whose cosine is the power factor. Therefore, for the small loss angles usually found in good dielectrics, the dissipation factor is equal to the power factor for all practical purposes.



Assembled Cavity

Outer Cylinder of Cavity Removed, Showing Internal Details

FIG. 3. 300–1000-Mc variable length re-entrant resonant cavity for measuring dielectric properties of materials.

Energy is coupled into the cavity by a loop as shown. By varying the size and position of this loop, the measured Q[8] of the cavity may be varied from less than 50 to about 2000. The detector consists of two type 1N21 crystals (Sylvania) side by side with ends of opposite polarity connected to the probe.

The fixed cavity has the limitation that the frequency at which the measurement of a given sample is made depends on the dielectric constant and dimensions of the sample. By a reasonable variation in the thickness of a sample, however, measurements may be made over the frequency band previously designated for this cavity.

2. The Variable Cavity

Because there may be cases where it is desirable to select the exact frequency at which a measurement is to be made, and also to obtain a single instrument useful over a wide band of frequencies, a variable length re-entrant cavity was designed for a frequency range of 300–1000 megacycles per second. Photographs of this instrument are shown in Fig. 3. Details of the cavity proper, which are not evident from these illustrations, appear in Fig. 4.

The length of the cavity is controlled by the position of a shorting plug. It might be emphasized that in this method variation in contact does not introduce an error in the measurements. The plug is driven by a threaded tube which is restrained from turning by a key on the extension

[8] Q is usually defined as the ratio of the amount of energy stored per cycle to the amount of energy dissipated per cycle. Q is thus the reciprocal of the dissipation factor.

of the center post. This tube is propelled by a nut to which the dial (see Fig. 3) for indicating the length of the cavity is attached. A vernier drive is also provided.

The separation of the electrodes is controlled in the same manner as in the fixed cavity. In fact, all other details of the cavity are similar to the fixed cavity except that the former is smaller in diameter and has smaller electrodes, as indicated in Fig. 4. This smaller diameter results in a lower Q. By adjusting the coupling, the measured Q may be caused to vary from less than 50 to about 1000. This maximum value of Q still gives good sensitivity in measuring the dissipation factor of low loss materials. The sample required is a small disk 0.480″ or less in diameter, and from 0.040″ to 0.125″ thick. The auxiliary equipment used with these cavities is described in reference 6.

III. THEORY

1. General Considerations in the Design of the Cavities

The well-known susceptance variation method[9,10] for the measurement of the dielectric constant and dissipation factor in the frequency range 10^4 to 10^8 cycles per second has a number of advantages. It combines a simple and rapid measuring technique with simple relations for rapid computation of results with an accuracy greater than that obtainable by most other methods. It has been found possible to apply this technique to measurements at higher frequencies. Expressions were derived for the dielectric constant and loss of samples, measured by placing them in the re-entrant end of a cavity and using a measuring procedure similar to that of the conventional susceptance variation technique. These are identical, except for an additional correction term, with those used in the usual application of this technique at radio frequencies, provided the dimensions of the sample are so restricted that it can be considered as a lumped capacitance.

If the sensitivity to low loss samples is to be satisfactory, the Q of the cavity must be high. The expression for the theoretical Q of such cavities is[11]

$$Q=4\pi(G/10)^{\frac{1}{2}}b\frac{\log_e b/a}{b/a+1}(f)^{\frac{1}{2}}\times 10^{-4}, \qquad (1)$$

where G is the conductivity of metal covering inside of cavity in mhos per centimeter (5.8×10^5 for copper), b is the inside radius of the outer conductor of the cavity in centimeters, a is the outside radius of inner conductor of cavity in centimeters, and f is the frequency in cycles per second. The optimum Q is obtained with a ratio of b/a of about 3.6. Examination of the curve[12] of Q *versus* the ratio b/a shows that this curve is fairly flat for values of b/a between 3 and 4, and falls off rapidly on either side of these values. In the frequency band being considered, a theo-

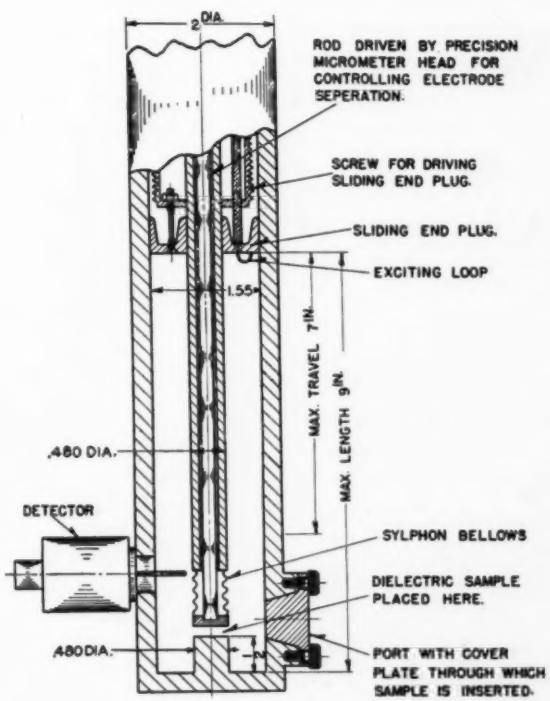


FIG. 4. Variable length re-entrant cavity.

[9] L. Hartshorn and W. H. Ward, "The measurement of the permittivity and power factor of dielectrics at frequencies from 10^4 to 10^8 cycles per second," J. Inst. Elec. Eng. **79**, 597–609 (1936).

[10] A.S.T.M. Committee D-9 "Tests for power factor and dielectric constant, susceptance variation method," *A.S.T.M. Standards on Electrical Insulating Materials* (The American Society for Testing Materials, Philadelphia, Pennsylvania, 1944), p. 341.

[11] Slater, *Microwave Transmission* (McGraw-Hill Book Company, Inc., New York).

[12] R. C. Miedke, "Q for unloaded concentric transmission lines," Electronics 139–140 (September 1943).

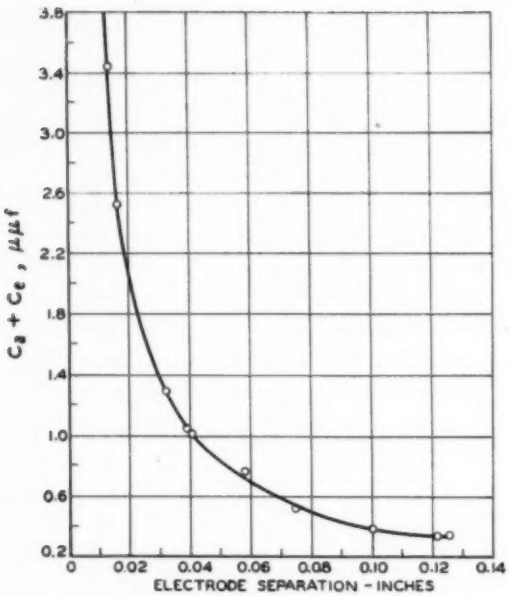


FIG. 5. Curve for the determination of $(C_{a_2}+C_{e_2})-(C_{a_1}+C_{e_1})$.

retical Q of about 4000 may be obtained for values of b as small as 2 cm. Because of energy absorbed by the probe and the effect of the coupling loop, the theoretical Q is never obtained. Measured Q of one-quarter to one-half of the theoretical Q may be obtained under practical operating conditions.

Having established the restrictions imposed to obtain satisfactory Q values, the next step is to determine the dimension of cavities, which resonate in the desired frequency band. The resonant frequency of the singly re-entrant cavity for the general case is treated by Hansen.[13] This case is so complex that it is at once apparent that it would be desirable to impose restrictions on the shape of the cavity which would result in simpler relations. E. U. Condon[14] gives an approximate treatment, which assumes that b is small compared with a quarter wave-length. He also shows that this treatment corresponds to applying the standard engineering form of transmission line theory. Therefore, this restriction will be applied to these cavities, and transmission line formulae used to calculate the resonant frequency or resonant length. The resonant frequency of a short circuited loss-free transmission line with a capacitor at one end is given by the equation[11]

$$Z_0 \tan\frac{2\pi fS}{c}=\frac{1}{2\pi fC}, \qquad (2)$$

where Z_0 is the characteristic impedance of coaxial part of cavity, S is the length of internal conductor, c is the velocity of light, and C is capacitance between the ends of the re-entrant posts in the cavity at resonance. The resonant length of the cavities, as calculated from this relation, checked the resonant length as determined experimentally within two percent.

Experiments with a singly re-entrant cavity indicated that the fringing of the field is excessive, especially for very short gaps. Therefore, the doubly re-entrant cavities of Figs. 2 and 4 were selected in which one of the re-entrant posts is much shorter than the other. On first consideration, it would seem that the re-entrant posts should be the same length, giving a symmetrical cavity which might be easier to analyze. However, the requirement of variable gap length would disturb this symmetry. It was observed that for small gap spacings (less than $0.02S$) and for the short post less than $0.25S$, the resonant frequency of the doubly re-entrant cavity varied

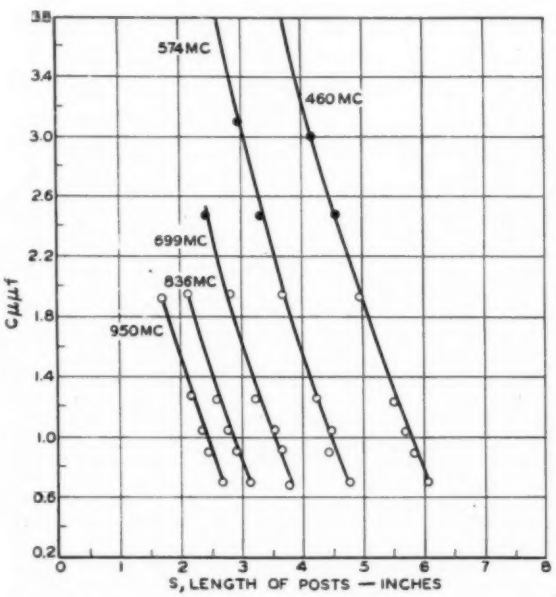


FIG. 6. Curves showing relation between geometric capacitance, C, between electrodes *versus* length of posts at constant resonant frequency.

[13] W. W. Hansen, "On the resonant frequency of closed concentric lines," J. App. Phys. **10**, 38–45 (1939).

[14] E. U. Condon, "Principles of Microwave Radio," Rev. Mod. Phys. **14**, 341–389 (1942).

very slightly from that of a singly re-entrant cavity of the same total length and same gap spacing. This observation confirms the experimental results obtained by Barrow and Mieher.[15]

The lowest mode of oscillation of this cavity corresponds to that of a conventional coaxial line short circuited at one end and capacitatively loaded at the other. It is usually designated $TM_{0,0,p}$ where p is less than unity. Since the diameter of our cavities has been restricted so that the resonant frequency could be calculated by transmission line methods, it is unlikely that higher modes could exist. Calculation of the cut-off frequency for the next higher mode verified this.

2. Expressions for the Dielectric Constant

By definition, the dielectric constant, e', of any dielectric sample is given by

$$e' = C_x / C_a, \tag{3}$$

where C_x is the capacitance of the sample, and C_a is its equivalent air capacitance. The capacitance of the sample, C_x, is given by the equation[6]

$$C_x = (C_{a2} + C_{e2}) - (C_{a1} + C_{e1}) + C_{a1} - \frac{dC}{dS}\Delta S. \tag{4}$$

Except for the last term, this is the same expression used in the conventional susceptance variation method.[10]

In applying this equation to the cavities, the more complicated case of the variable length cavity will be considered first. The capacitance of the electrode arrangement of two parallel cylinder faces is determined at lower frequencies by calibration against a standard capacitor in the case of the conventional susceptance variation circuit. In the frequency band being considered here, neither a standard capacitor nor the technique of comparing it to an unknown is available. Therefore, the calibration was determined experimentally, using two materials having considerably different values of dielectric constant, polystyrene and Insanol (mica-filled lead borate). The dielectric properties of samples of these materials were known from measurements made at higher and lower frequencies by other independent methods.

From the measurement of these materials of known properties, the calibration curve of Fig. 5 was obtained. This corresponds to the correction for fringing usually applied in the case of the conventional susceptance variation circuit. By obtaining the relation of the capacitance between electrodes to the length of the posts at constant resonant frequency, the curves of Fig. 6 were

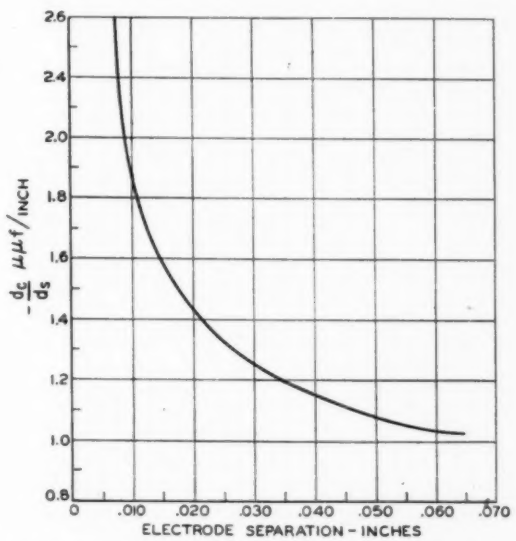


FIG. 7. Curve for the determination of dC/dS for any electrode separation.

obtained. The curve of Fig. 7 was then obtained by plotting the slope of these curves *versus* electrode separation. Multiplying the ordinate of this curve by ΔS; the change in the gap spacing to retune the cavity to resonance upon removal of the sample, gives the last term of Eq. (4). This term is the correction for the small change in length of the coaxial part of the cavity which necessarily takes place when one of the posts is extended inward to diminish the gap.[16]

Now Eqs. (3) and (4) may be solved as follows to obtain the value of the dielectric constant of any material. The quantities $(C_{a2}+C_{e2})-(C_{a1}+C_{e1})$, and $-dC/dS$ are read from the curves in Figs. 5 and 7, respectively. The value of $C_{a2}+C_{e2}$

[15] W. L. Barrow and W. W. Mieher, "Natural Oscillations of Electrical Cavity Resonators," Proc. I.R.E. **28**, 184–191 (April 1940).

[16] The reader who may wish to perform the experiments described should refer to the thesis "A Variable Length Re-entrant Resonant Cavity for the Continuous Measurement of Dielectric Properties from 300–1000 Mc," University of Pittsburgh, 1946, for a detailed discussion of the method used to obtain Figs. 5, 6, 7.

is read from the ordinate of Fig. 5, corresponding to the value of electrode separation with the sample removed. Similarly, the value of $C_{a1}+C_{e1}$ is found from the separation with the sample in place.

In the case of the fixed cavity, samples of various thicknesses, whose dielectric constants were accurately known, were measured over the frequency range of 400–600 megacycles per second. Using the computed theoretical values

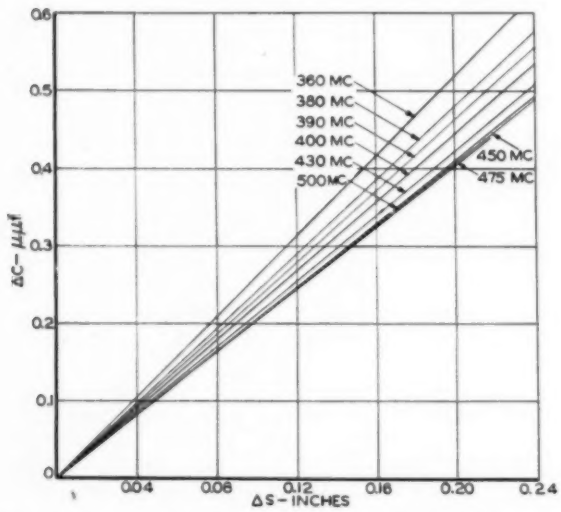


FIG. 8. Correction curve for the 400–600-Mc fixed cavity.

for the correction term, $(dC/dS)\Delta S$, the measurements all checked the known values within one percent. This indicated that fringing could be neglected in the range of frequency in which this equipment operates, and that the term $(dC/dS)\Delta S$ gives the total required correction. Then Eq. (4) may be written

$$C_x = C_{a2} - \frac{dC}{dS}\Delta S. \tag{5}$$

Replacing the term $-(dC/dS)\Delta S$ by its value, ΔC, as obtained from the curves of Fig. 8 and writing for C_{a2} its value for the case of electrodes 1″ in diameter, the following expression is obtained

$$C_x = \frac{0.1764}{\text{Gap spacing (inches) at resonance with sample removed}} + \Delta C \text{ u.u.f.} \tag{6}$$

The simplicity of Eqs. (3) and (6) permits very rapid calculation of the dielectric constant.

3. Expressions for the Dissipation Factor

An expression for the dissipation factor of a sample placed between the ends of the posts of a re-entrant cavity was derived in an earlier paper.[6] It is identical with the equation used for the "Change of Voltage Procedure" in the well-known susceptance variation method used at radiofrequencies.[10] The final result is given by the following expression:

$$\text{Tan}\,\delta_x = (\Delta C_2/2C_x)(E_2-E_1)/E_1, \tag{7}$$

where E_1 is the voltage at resonance with the sample in place, and E_2 is the voltage at resonance with the sample removed. The term $\Delta C_2/2C_x$ should be recognized as the dissipation factor of the empty cavity. ΔC_2 is the change in capacitance necessary to tune from one half power level of resonance to the other.

IV. PERFORMANCE

The dielectric constant and dissipation factor of a number of materials over a frequency range of 60 cycles per second to 10,000 megacycles per second are shown in Figs. 9 to 11. These curves are given to show the agreement in the results obtained using the instruments and techniques described with values obtained by independent methods. Symbols as listed on each curve are used to identify the method used to obtain a particular value.

Figure 9 shows the dielectric properties of polystyrene. The properties of this material are known to be practically constant over the frequency range shown on this curve. Therefore, the measurement of this material provides a

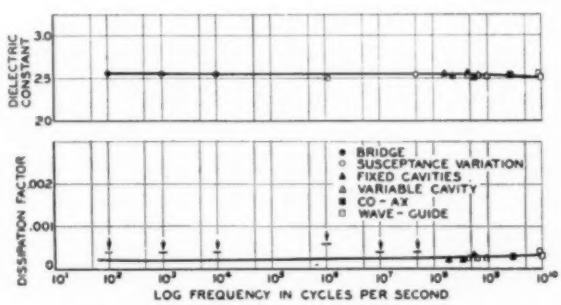


FIG. 9. Dielectric properties of polystyrene.

convenient check as to whether the measurements made by the method described agree with measurements made by independent means at higher and lower frequencies. Inspection of the curve for the dielectric constant shows that the values for the frequency decade 10^8 to 10^9 cycles per second lie along a curve already well defined by values determined by independent methods. Likewise, Fig. 10 shows that the values for the dielectric constant of insanol, which has a value of about 8, in the frequency decade in which we are interested do not depart from the curve. This shows that the accuracy does not vary with change in the value of the dielectric constant. The accuracy of the value of dielectric constant, as indicated by these curves, and other unreported measurements, is about ± 1 percent.

The value of the dielectric constant of most solid materials is usually between two and ten. In the case of those unusual materials whose dielectric constant is considerably greater than ten, the samples should be much smaller than the electrodes.[17] A sample with a dielectric constant much greater than ten, which occupies the entire space between the electrodes, shortens

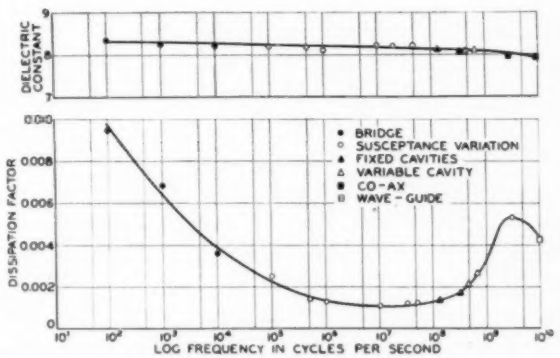


FIG. 10. Dielectric properties of insanol (mica-filled lead borate).

the effective length of the cavity to such an extent that even when the physical length of the cavity has been made a minimum, it is not possible to resonate the cavity at frequencies in the upper portion of the band for which the instrument was designed.

The error in the dielectric constant caused by variation in voltage across the face of the electrodes was investigated and the proper correction

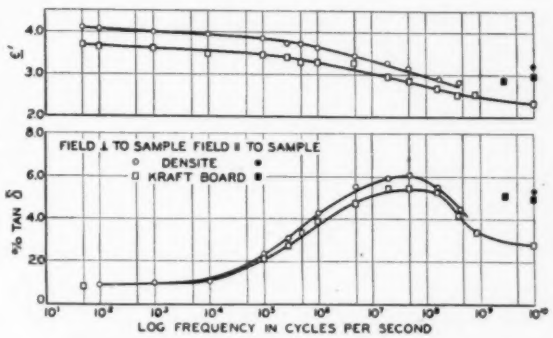


FIG. 11. Dielectric properties of kraft board dried at 105°C for 16 hours.

made in all cases where the error was not negligible. When materials having dielectric constants greater than five are measured at frequencies over 500 megacycles in the cavities described, a correction should be made for the variation in voltage across the electrode faces.[18]

The curve for the dissipation factor, Fig. 9, shows that these values are in line with the values at higher frequencies obtained by wave guide measurements. At lower frequencies, the precision of the apparatus is ± 0.0002 so that all that may be stated with certainty is that the dissipation factor is less than 0.0004. Further examination of the curves of dissipation factor, Figs. 9 and 10, shows that the accuracy of these values, as obtained by both the fixed and variable cavities, is ± 0.00005 for small values of $\tan\delta$.

In spite of the fact that certain approximations and empirical calibrations have been made to keep the measuring procedure and calculations nearly as simple as those of the conventional

[17] When the sample is smaller than the electrodes, Eq. (4) is applied to determine the over-all dielectric properties of the space between the electrodes (with the sample in place between the electrodes), and then standard equations for dielectrics in parallel (sample and air) are applied to determine the properties of the sample.

[18] The capacitance and, therefore, e' of a sample placed between electrodes varies as: $e_c' = e_m'(E_{max}/E_{Av})$ where e_c' is the corrected dielectric constant and e_m' is the dielectric constant as calculated by Eq. (3). The average voltage is given by the expression $E_{Av} = E_{max}[1-(\pi^2 e' r^2 f^2/2c^2)]$ where r is the radius of the sample, and the remaining symbols have been defined previously. This last equation may be derived by using Eqs. (7'12), (7'6), and (7'2) of reference 12, expanding J_0 in an infinite series, and taking the average. Because only the first two terms of the expression are given, r must be less than $\lambda/8$, where λ is the wavelength in air.

susceptance variation method, the accuracy of the values of the dielectric constant is very nearly the same as that obtained through the use of this technique at lower frequencies. The accuracy of the values of low dissipation factors is much greater than that obtained by any of the methods used at lower frequencies. This is due to the fact that very high values of Q are obtained with the resonant cavity. This property of the resonant cavity may be taken advantage of through suitable techniques to obtain the dissipation factor of low loss dielectrics to very great precision.[19]

The curves of Fig. 11 are included because they illustrate the usefulness of apparatus operating in this frequency range, and point out some interesting properties of dielectrics. These materials have a dispersion region. The dielectric constant decreases as the dissipation factor goes through a maximum between 10^7 and 10^8 cycles per second. The apparatus described made possible an accurate determination of these curves throughout the dispersion region. Few laboratories have an apparatus for the measurement of solids operating in the frequency band of 10^8 to 10^9 cycles per second. These materials have another interesting property that should be mentioned in passing. The direction of the electric field in the sample makes a considerable difference in the values of the dielectric properties. When this material is placed in a wave guide with the same face against the end of the guide which was in contact with the electrode, the direction of the field in the sample is changed by 90°. Therefore, when changing from an instrument with electrodes to a wave guide, consideration should be given to the direction of the field in the sample in the case of materials which may be non-isotropic.

[19] William R. MacLean, "A microwave dielectric loss measuring technique," J. App. Phys. **17**, 558–566 (1946).

V. CONCLUSION

Although the calibration of the apparatus described in this paper was accomplished by empirical methods, it has many advantages to recommend its use over methods which are mathematically more rigorous.

The sensitivity of these re-entrant cavities is so great that very small disk-shaped samples can be used. This makes it possible to sample directly the material to be measured. Thus, a piece of dielectric material may be removed from a cable or other apparatus and tested. When large samples of intricate shapes are required, as in the case of coaxial transmission line methods, especially at frequencies near the lower end of the frequency band (100–1000 Mc) considered in this paper, samples molded or cast in the laboratory may not be representative of a commercial grade of the same material. The simple disk shape of the samples required for this method makes their preparation easy.

Measurements may be made rapidly, and the calculation of results is extremely simple. The accuracy obtained from this method is usually better than that of other methods operating in this frequency range.

VI. ACKNOWLEDGMENT

The author wishes to make grateful acknowledgment to Drs. L. J. Berberich and T. W. Dakin of the Westinghouse Research Laboratories, and to Dr. E. W. Montroll of the University of Pittsburgh for the many helpful suggestions given during the course of this work.

A Variable Capacitor for Measurements of Pressure and Mechanical Displacements; A Theoretical Analysis and Its Experimental Evaluation*

JOHN C. LILLY, VICTOR LEGALLAIS, AND RUTH CHERRY
E. R. Johnson Foundation for Medical Physics, University of Pennsylvania, Philadelphia, Pennsylvania
(Received December 13, 1946)

A variable capacitor is described for measuring (1) small displacements, (2) small volume changes, and (3) pressure differences. The capacitor consists of a deflectable diaphragm and a fixed electrode. The diaphragm is metallic, plane-parallel, clamped at the edges, and at ground potential; the electrode, at an a.c. potential, has a plane surface parallel to the undeflected plate across an air gap. For use in displacement measurements, the diaphragm's center is deflected by a point contact from a mechanical link to the observed system, or by a uniform pressure load from a fluid link to the system. The fluid link is used also when measuring volume changes and pressure differences. The plate deflection results in a change in the air gap, and thus generates a capacitance signal. This signal is measured by electrical methods.

A theoretical analysis of this variable capacitor is presented; sensitivity and alinearity factors for the three uses of the device are derived. The experimental performance shows reasonably satisfactory agreement with the derived theory. The displacement of the plate's center was measured with an interferometric method, using a yellow He line as a standard of reference; the applied pressure, with a liquid manometer; and the capacitance signal, with a standard capacitor substitution procedure. The gauge can be used so as to give quantitative electrical indications of displacement, volume change, or pressure difference; or can be used as a null indicator device in which an unknown pressure is balanced against a known one on opposite sides of the diaphragm. In order to achieve large volume an displacement sensitivities, small air gaps (5.10^{-4} cm) are employed. Details of a construction method to assure small values are presented.

INTRODUCTION

TO anyone who has attempted the design or the construction of a diaphragm capacitance gauge for precise pressure or volume-change measurements, the lack of an adequate theory to handle the many variables encountered is rather discouraging. The important variables are the diaphragm diameter, thickness, and material; the air gap, the electrode diameter, the applied pressure, and the capacitance signal. The magnitude of the pressure signal and the desired capacitance signal are known. But what about the values to be assigned the first five variables in order to achieve the desired pressure-capacitance relation? It was an attempt to answer this question that led to this analysis.

Though the theory presented here was evolved in response to a need in the biological field, the treatment and the results are sufficiently general to be of interest to the physicist, the chemist, or the engineer who may want a gauge for other applications. For the benefit of those who may apply the gauge to biological systems, and for the benefit of those interested in why biological applications are as strict in their requirements as the physical ones, the peculiarities of the living systems are presented and discussed in relation to the capacitor gauge design. For those interested only in gauge design, this discussion on biological problems may be neglected, and a start made on the later, theoretical derivation section.

THE GAUGE DESIGN PROBLEM IN BIOLOGICAL APPLICATIONS

The pressure gauge design problems encountered in the biological field are different from those in other fields of research. Certain limitations are placed on the pick-up unit design by the special properties of animals. In general, the pressure systems in animals are buried within the body, loosely supported, of variable distensibility, in constant motion, and are fragile in construction. These structures react locally to probing, puncturing, and tearing by bleeding and by constricting. Blood in the vessels may clot on inserted objects, and thus plug pressure taps. The whole animal also reacts by moving and by markedly changing the blood pressure, the heart rate, and other variables one may be measuring. Anesthetics can be used to abolish many of these undesired effects; but these drugs can exert other actions which are also unwanted.

* The work described in this paper was done under a contract recommended by the Committee on Medical Research, between the Office of Scientific Research and Development and the University of Pennsylvania.

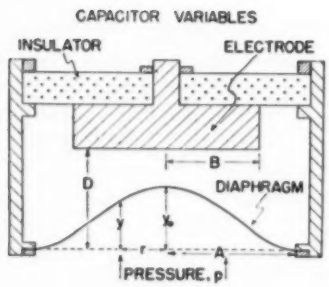


FIG. 1. Diagram of a diaphragm capacitor. The variables used in developing the capacitor theory are illustrated here; capital letters refer to fixed geometrical properties, small letters to variables generated by a displaced diaphragm. Tables I and II give the definitions and the units of all variables. The diaphragm is shown in an approach deflection; in a recession deflection the diaphragm would bulge downward. Both the air gap and the center deflection are greatly magnified in relation to the diameters of the diaphragm and of the electrode. The form of the deflected diaphragm is that of a stiff, edge-clamped plate; Figs. 2 and 3 give quantitative details of the two extreme idealized diaphragm forms assumed in the theory.

The pressure phenomena occurring in animals have both a static level and superimposed rapid pulsations. It is important to record these phenomena faithfully.

Because of these special properties of the animal, some form of transmission of the pressure signal from the interior to the exterior is necessary. At the present time, tubing containing liquid usually conducts the signal; ultimately, electrical transmission may be substituted. To avoid undesired local stimuli, the tubing is small.

From these considerations, some requirements for a pressure gauge pick-up can be deduced. (1) The pick-up must reproduce static level signals. (2) For the faithful reproduction of rapidly changing pressures transmitted through small bore tubing, and to minimize the length of the tube bore exposed to coagulable blood, the volume change of the unit with respect to an applied pressure must be minimal. (3) The pick-up size and the weight should be as small as possible. (4) The connection between the pick-up unit and the recorder must be long and very flexible. (5) The pick-up and its leads should be insensitive to vibration and to changes in position.

Many types of diaphragm pressure gauges have been used in recording static and kinetic pressures in animals.[1–8] The ones with the greatest promise of meeting the above requirements are the photoelectric[6,7] and the electrical capacitor[5] types. The latter type is the one analyzed here.

With the above conditions in mind, it was considered important to examine the theoretical basis of the pick-up unit of a diaphragm-type capacitor pressure gauge previously applied to the recording of pressure in the vascular system.[8] The examination, presented below, was undertaken in order (1) to find out what factors were of importance in an optimal design for a pick-up unit for a minimal volume change and maximal capacitance signal in response to a given applied pressure; (2) to facilitate making design compromises in the presence of construction difficulties in special units; (3) to find and to evaluate those factors in the design which determine the alinearity and the sensitivity of the gauge; and (4) to explain and thus make allowances for the observed differences in gauge

TABLE I. Measured and calculated variables for a given gauge.

I. Measured variables	Units	Method
1. Electrode radius	B cm	
2. Diaphragm radius	A cm	By micrometer
3. Diaphragm thickness	T cm	
4. Applied pressure	p dynes/cm²	By liquid manometer
5. Central point deflection, at p	y_0 cm	By interferometer
6. Displaced volume, at y_0	v cm³	By calibrated capillary
7. Gauge capacitance, at $y_0=0$	C_0 cm	By substitution for standard condenser in an electrical circuit
8. Capacitance signal, at y_0	c cm	
II. Calculated variables, as derived from measured ones		
9. Air gap, at $y_0=0$	D cm	$D=B^2/4C_0$
10. Displaced volume, at y_0	v cm³	$v=(y_0)(\pi/k)A^2$
11. Relative deflection	w (0)	$w=y_0/D$
12. Radius ratio, squared	N (0)	$N=(B/A)^2$
13. Relative capacitance signal	s (0)	$s=c/A^2/4D=(c/C_0)N$
14. Deflection constant, stiff	K cm²/dyne	$K=(y_0/p)(T^3/A^4)$ $K=\frac{3}{16}(1-t^2)/Y$ (t=Poisson's ratio; Y=Young's modulus)
15. Deflection constant, flexible	$1/G$ cm/dyne	$1/G=(y_0/p)/A^2$
16. Volume constant, stiff	k (0)	$k=y_0(\pi A^2)/v=3$
17. Volume constant, flexible	k (0)	$k=y_0(\pi A^2)/v=2$
18.	H (0)	$H=(\frac{1}{3}-N)$
19.	u (0)	$u=w^{\frac{1}{3}}$
III. Calculated variables, as related in theory: see equations in Table III		

[1] O. Franck, Zeits. f. Biol. **82**, 49 (1925).
[2] C. J. Wiggers, J. Lab. and Clin. Med. **10**, 54 (1924).
[3] W. F. Hamilton, Am. J. Physiol. **107**, 427 (1934).
[4] A. Hampel, Pflunger's Arch. Ges. Physiol. **244** (2), 171 (1940).
[5] W. G. Kubicek, Rev. Sci. Inst. **12**, 101 (1941).
[6] H. Rein, Arch. f. d. Ges. Physiol. **243**, 329 (1940).
[7] W. E. Gilson, Science **95**, 514 (1942).
[8] J. C. Lilly, Rev. Sci. Inst. **13**, 34 (1942).
[8a] F. Buchthal and E. Warburg, Acta Physiol. Scandinavica **55**, 55–70 (1943).

TABLE II. Sensitivity and alinearity factors for a constant B/A and a given w.

I. Sensitivity factors	Secant form	Tangent form	Initial value, $y_0 \to 0$
20. Relative displacement	$m=s/w$	$n=(\partial s/\partial w)$	$M=(1-H^2)/k$
21. Displacement	$c/y_0=(m/4)(A/D)^2$	$\partial c/\partial y_0=(n/4)(A/D)^2$	$S_y=(M/4)(A/D)^2$
22. Volume	$c/v=km/4\pi D^2$	$\partial c/\partial v=kn/4\pi D^2$	$S_v=kM/4\pi D^2$
23. Pressure	$c/p=mKA^6/4T^2D^2$	$\partial c/\partial p=nKA^6/4T^2D^2$	$S_p=MKA^6/4T^2D^2$
23a. Pressure	$c/p=(c/y_0)(y_0/p)$	$\partial c/\partial p=(\partial c/\partial y_0)(\partial y_0/\partial p)$	$S_p=(M/4)(A/D)^2(y_0/p)$
II. Alinearity factors			
24. Relative displacement	$a=m/M$	$b=n/M$	$a=b=1.00$
25. Displacement	$a_d=(c/y_0)/S_y=a$	$b_d=(\partial c/\partial y_0)/S_y=b$	$a_d=b_d=1.00$
26. Volume	$a_v=(c/v)/S_v=a$	$b_v=(\partial c/\partial v)/S_v=b$	$a_v=b_v=1.00$
27. Pressure	$a_p=(c/p)/S_p=a$	$b_p=(\partial c/\partial v)/S_p=b$	$a_p=b_p=1.00$

TABLE III. Theoretical equations relating s, w, and B/A.*

Stiff plate (Fig. 2)		Flexible membrane (Fig. 3)	
	Approach case		
(2.0)	$s=(1/u)\tanh^{-1}[uN/(1-wH)]-N$	(2.1)	$s=(1/w)\ln[(1-wH)/(1-w)]-N$
(3.0)	$s=[(\frac{1}{3})(1-H^3)w+(\frac{1}{5})(1-H^5)w^2+\cdots]$	(3.1)	$s=[(\frac{1}{2})(1-H^2)w+(\frac{1}{3})(1-H^3)w^2+\cdots]$
(4.0)	$n=(\frac{1}{2})[1/(1-w)-H^3/(1-wH^2)-m]$	(4.1)	$n=1/(1-w)-H^2/(1-wH)-m$
	Recession case		
(5.0)	$s=(1/u)\tan^{-1}[uN/(1+wH)]-N$	(5.1)	$s=(1/w)\ln[(1+w)/(1+wH)]-N$
(6.0)	$s=[-(\frac{1}{3})(1-H^3)w+(\frac{1}{5})(1-H^5)w^2-\cdots]$	(6.1)	$s=[-(\frac{1}{2})(1-H^2)w+(\frac{1}{3})(1-H^3)w^3-\cdots]$
(7.0)	$n=(\frac{1}{2})[1/(1+w)-H^3/(1+wH^2)-m]$	(7.1)	$n=1/(1+w)-H^2/(1+wH)-m$

* Note that (Table I): $u^2=w$; $N=(B/A)^2$; $H=(1-N)$.
The variables are defined in Figs. 1, 2, and 3, and in Table I. (1) Fundamental integral: $s=[(4D/A^2)\int_0^B(rdr)/2(D\pm y)]-N$. Solutions: [from (1) and equations of Figs. 2 and 3].

performance in the case in which the diaphragm is deflected away from the electrode in contrast to the case of diaphragm-electrode approach.

In addition to the measurement of pressure, a diaphragm-type capacitor has desirable characteristics for use as an ultramicrometer (displacement) and as a dilatometer (volume change). In the theory below, the condenser characteristics for these three uses are most easily discussed in the order: displacement, volume, and lastly, pressure. (In the theory, displacements are considered only as measured by means of a fluid link; in practice, point contact mechanical links can be used, but this theory does not apply in this case, even with a stiff plate.)

The following sections give a summary of a derivation of a diaphragm capacitor theory, and some applications of the theory in the design and in the testing of diaphragm capacitors. Later sections deal with the results of an experimental check of the theory, a consideration of electrical circuits for recording with the gauge, and a method used in the construction of some pick-up units.

DERIVATION OF A THEORY OF A DIAPHRAGM-TYPE CONDENSER

A. Configuration of the Capacitor

In deriving a theory of a diaphragm-type capacitor an expression is first developed which relates the displacement of the diaphragm to the resultant capacitance signal, for the geometrical arrangement shown in Fig. 1. In the following discussion, variables generated by a diaphragm deflection are represented by small letters, and the constants of a given gauge, by capital letters (see also Tables I and II).

The diaphragm is a circular plane-parallel metal plate or membrane clamped near the periphery and is assumed to be deflected by a uniform pressure load on one side or the other. The effective radius of this diaphragm, A, is that of the periphery at the inside edge of the clamp.

The electrode is a circular metal structure with a plane surface placed parallel to the undeflected plate, across an interposed air gap of a thickness D. The effective electrode radius, B,

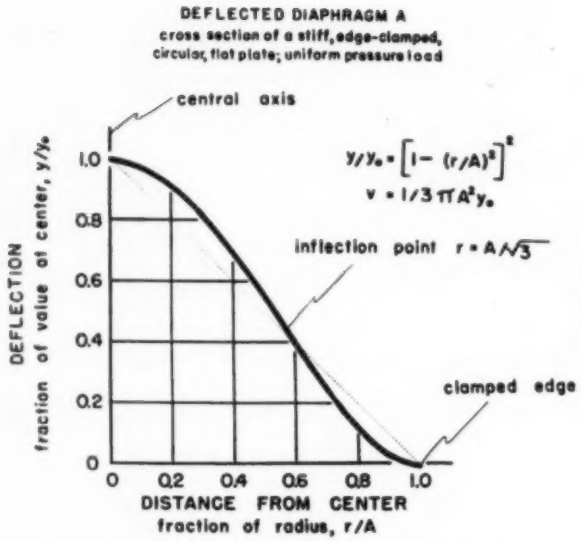


FIG. 2. The form taken by a stiff plate deflected by a uniform pressure. A radial section is cut through the diaphragm, from the axis of rotation at the left, to the clamped edge at the right. This method of showing the form greatly magnifies the amount of the deflection in relation to the radius by a factor of about one to ten thousand. The equation of this curve is derived on the assumptions that the absolute deflection is less than about 1/50 the thickness of the diaphragm, that the thickness is less than about 1/50 of the radius, and that the edge is uniformly and rigidly clamped, without exerting a radial tension in the undeflected position of the diaphragm (no initial stretching). All strains are assumed to be within the proportional elastic limit. It is further assumed that, in the undeflected state, the diaphragm is a plane, parallel, homogeneous, and isotropic structure. The equation of this curve is given in the figure; the variables are defined by Fig. 1. The volume swept out by the diaphragm as it is deflected, v, was found by integration of the equation for the curve; the resulting expression gives the dimensions of a cone of equivalent volume.

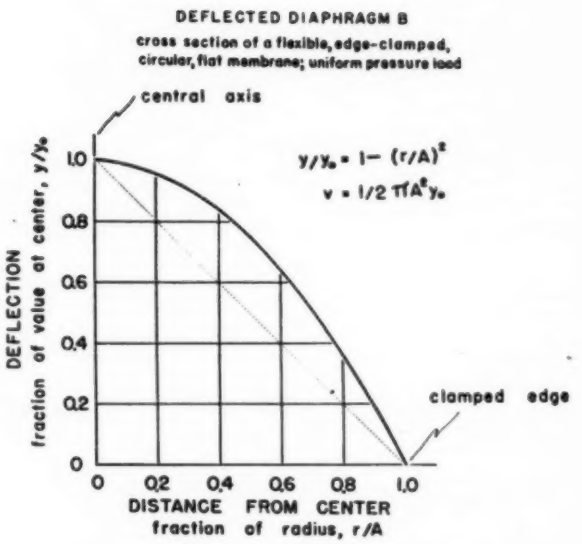


FIG. 3. The form taken by a flexible membrane deflected by a uniform pressure. The general statements made under Fig. 2 apply to this figure. The same limiting assumptions apply to this equation as to the stiff-plate one, except that it is assumed that the membrane is subjected to a uniform tension of $G/4$ cm/dyne (item 15, Table II) in all directions in its own plane in the undeflected state. In this diaphragm the inherent stiffness is assumed to be negligible, and hence the clamped periphery bends with a practically infinitesimal radius of curvature. The volume swept out by the diaphragm as it is deflected, v, equals that of a cylinder of a height equal to one-half the central deflection and a base of an area equal to that of the undeflected diaphragm.

is that of the plane surface presented to the air gap.

The diaphragm and the electrode have a common axis passing through their centers, normal to their facing surfaces (Fig. 1).

Since the ratio of the electrode periphery to the air-gap thickness, $2\pi B/D$, is usually of the order of magnitude of 750–7500; edge effects are considered to be negligible and are omitted in the following analysis.

B. Derivation Summary

An equation defining the relation of the relative change in capacitance of this capacitor to the deflection of the diaphragm (Fig. 1) can be derived in the following manner:

1. Consider an elemental capacitor consisting of an annulus of infinitesimal width, lying in the surface of the electrode and concentric with the central axis, and of a similar strip directly across the air gap, lying in the surface of the diaphragm. Let both strips have a radius of r, defined as the distance along a normal from the annulus to the axis. Assume that, as the plate is moved by a pressure load, the capacitance of this elemental capacitor, dc, (in centimeter units) varies inversely with the distance between the strips, and varies directly with the area of one strip ($2\pi r dr$) (Eq. (1), Table III, note).

In the usual case, the deflection of the diaphragm center is only about 1/500 to 1/5000 of the diaphragm radius; therefore it can be shown that the above assumption is justified.

2. Consider the annular element in the surface of the diaphragm in the deflected position. The displacement of this element from the undeflected position is taken as y, and that of the diaphragm center as y_0. The relation between the radius of the annulus, r, and y, should be con-

sidered for at least two extreme cases, the thin, stiff plate and the thin, flexible membrane.[9]

The equation relating y, y_0, r, and A for the stiff plate is given in Fig. 2; and that for the flexible membrane in Fig. 3. (These equations are taken from Love's *The Mathematical Theory of Elasticity*, and reference 9.) The equations are exact only for values of y_0/A which are very small compared with unity. Other assumptions and limits for these equations are given in the Figs. 2 and 3 captions.

3. By setting up the differential equation implied in assumption 1 above, substituting the expression for assumption 2, for the two diaphragm types and integrating the result, the equations of Table III can be derived. These equations, in Table III, state the relation between the relative capacitance signal, s, and the relative change in the air gap, w (equals y_0/D). In these expressions the parameter is the electrode-diaphragm radius ratio, B/A. All of the deflections of the diaphragm toward the electrode are treated as the "approach case," and deflections away from the electrode as the "recession case."[10] The equations of these two cases differ fundamentally only in the algebraic sign of the central diaphragm displacement term, y_0, which is positive in recession and negative in approach.

For calculation purposes, the arc tangent, the arc hyperbolic tangent, and the logarithmic forms of the solution were used for numerical values of y_0/D greater than 0.10, and the series forms for the lesser values (0.00 to 0.10).

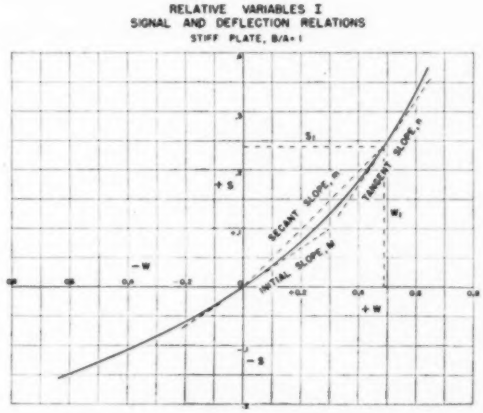


FIG. 4. Relation between the relative displacement and the relative capacitance change for a stiff plate. This curve is for an electrode and a stiff plate diaphragm (Fig. 2) of equal radii. For any condenser with a given radius ratio, B/A (Fig. 1), the fractional, or relative, capacitance change, s, is defined as the capacitance change, c, divided by the undeflected state capacitance, C_0, of a condenser with a B/A of unity and the same initial air gap, D (Fig. 1 and item 13, Table I). The fractional air-gap change or the relative displacement, w, is defined as the central diaphragm deflection, y_0, divided by the air gap, D, in the undeflected state (Fig. 1). This figure serves to define the relative displacement sensitivities, M, m, and n in terms of the slopes of the s and w curve. The secant m, and the tangent n, are both defined at any point on the curve, s_1 and w_1; the initial slope, M, is that at the origin, $s=w=0$. The sole dependence of the initial slope, M, on the radius ratio, B/A, is shown in the next figures. Both m and n approach M as the deflection, w, becomes infinitesimal. For the assumptions and the theory of this curve see text. It is to be noted that the variables used in this and in most of the rest of the figures have zero dimensions.

[9] I. B. Crandall, *Theory of Vibrating Systems and Sound* (D. Van Nostrand Company, New York, 1927).

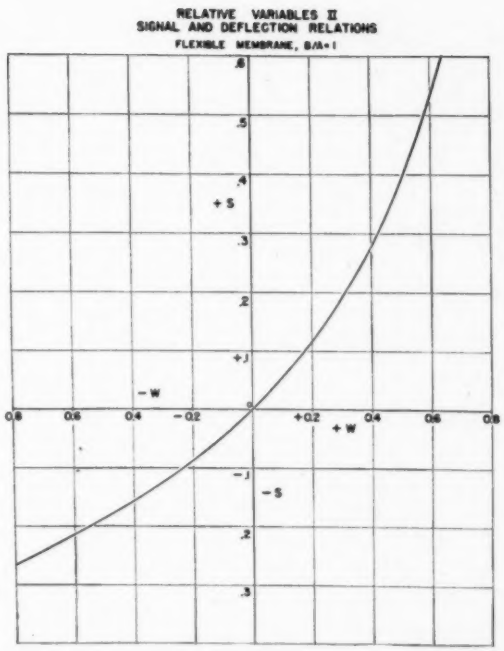


FIG. 5. Relation between the relative displacement and the relative capacitance change for a flexible membrane. The figure is the same as Fig. 4 with the exception that the diaphragm here is a flexible membrane (Fig. 3) instead of a stiff plate. The definitions of the variables are given under Fig. 4; the theory is developed in the text.

USE OF THE THEORY IN THE DESIGN OF PICK-UP UNITS

A. Factors Influencing the Sensitivity

In general the final objective in a problem of the design of a pick-up unit is reached when the unit produces a capacitance signal of the mag-

[10] Roess gives a stiff plate approach case equation which is equivalent to the inverse hyperbolic tangent function in Table III. [L. C. Roess, Rev. Sci. Inst. 11, 183 (1940).]

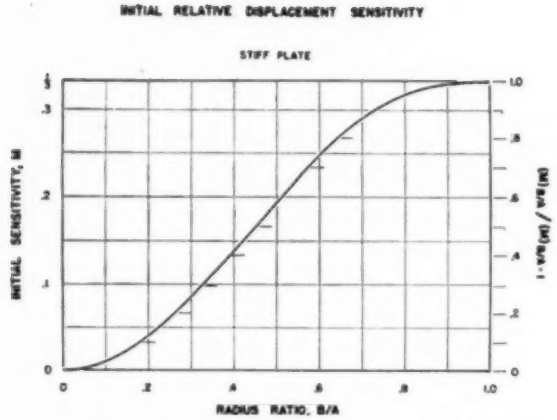


FIG. 6. The dependence of the initial relative displacement sensitivity on the radius ratio for a stiff plate. As is shown in the theory in the text, the initial relative displacement sensitivity, M (Fig. 4), is a function solely of the radius ratio, B/A (Fig. 1). The right-hand scale shows the sensitivity at a given B/A as a fraction of that at the maximum value, at B/A of unity. The initial absolute sensitivity, defined as the absolute capacitance change, c, divided by the absolute central deflection, y_0 (Fig. 1), can be calculated from these values of M, the electrode radius B, the diaphragm radius A, and the initial air gap, D (see text). The ratios of the secant, m, and the tangent, n, relative sensitivities (Fig. 4) to M for a number of radius ratios is shown in later figures.

nitude required by a chosen electrical circuit in response to an applied displacement, volume, or pressure change of a size determined by the system in which measurements are to be made. In other words, the designer's problem is to achieve a definite gauge sensitivity. In most biological applications to pressure recording, as was shown above, an additional part of the final objective is to have the volume of the gauge change a minimal amount. In some applications, a minimal alinearity of the calibration curve is also desirable.

Two useful types of sensitivity factors can be defined for diaphragm gauges; the absolute and the relative types. The absolute factors are defined in terms of the variables measured while the gauge is in use, which are the displacement, the volume, or the applied pressure, and the resulting capacitance signal. The relative sensitivity factors are expressed in terms of dimensionless quantities, such as those used in the theoretical equations developed above. After the gauge constants are known, the absolute sensitivity factors can be calculated from the relative sensitivity ones. This calculation will be dealt with in a later section.

B. Relative Sensitivity

Figures 4 and 5 show the graphs of the function relating the relative capacitance signal, s, to the relative displacement factor, y_0/D, for a radius ratio, B/A, of 1.00 (Tables I and III). As the plate or membrane deflects an infinitesimal amount in either direction (y_0/D approaching zero), the slope of this curve has a definite value. From inspection of the series forms of the equations in Table III, it can be seen that this slope at the origin is a function only of the radius ratio, B/A; this slope is called the "initial relative displacement sensitivity," M, for a given value of the radius ratio, B/A. Figures 6 and 7 show the relation between M and B/A for the two extreme diaphragm types.

As the relative displacement, w (equal to y_0/D, Table I), increases beyond an infinitesimal value, the relative sensitivity changes from the initial value, M. Two useful types of relative sensitivity factor can be defined for the rest of the curve; these two types are illustrated in Fig. 4. The slope of a straight line from any point on the

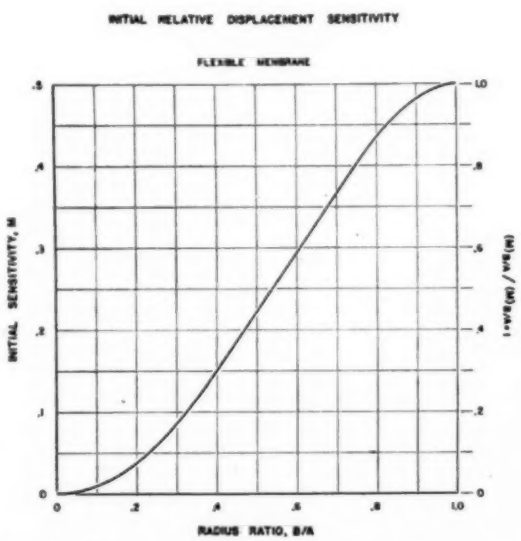


FIG. 7. The initial relative displacement sensitivity and the radius ratio for a flexible membrane. The statements under Fig. 6 apply also to this figure. It is to be noted that the maximum value of M at a B/A of unity for the flexible membrane is one-half, and that for the stiff plate (Fig. 6) is one-third. From the expression for the initial absolute volume sensitivity, S_v, (item 22, Table II), and these values of M and the volume constant, k (16 and 17, Table I), it can be shown that the initial absolute volume sensitivity for the stiff plate is equal to that of the flexible membrane considering all other geometrical values as identical, and B/A as unity.

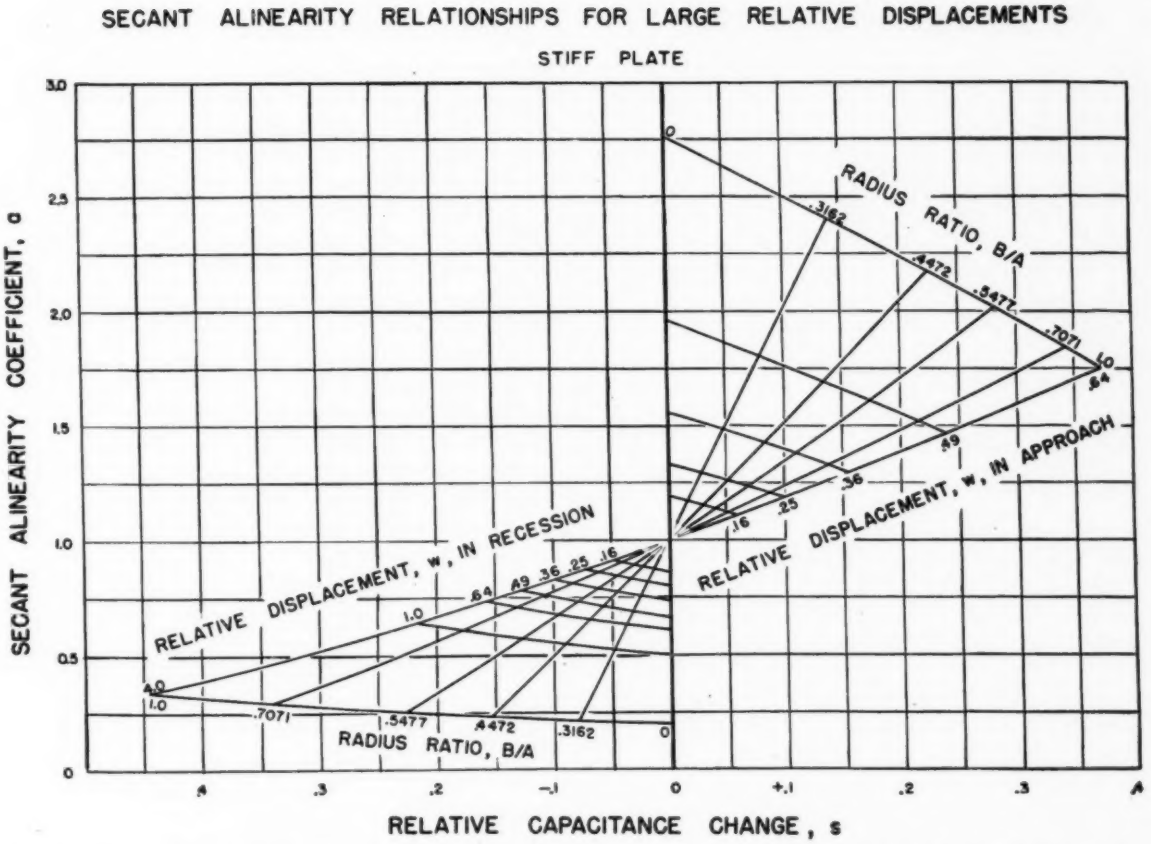


FIG. 8. The secant alinearity coefficient and the relative capacitance signal relations for a stiff plate for large relative deflections. Some of the variables used in this graph are defined under Fig. 4, all variables are defined in Tables I and II. The secant alinearity coefficient, a, is the ratio of the secant relative displacement sensitivity, m, to the initial displacement sensitivity, M. This figure shows the dependence of a on B/A, s, and w for a stiff plate. A point to be noted is the large deviations from linearity in the approach condition as contrasted with the smaller deviations in recession, for corresponding values of w.

curve to the origin is called the "secant relative sensitivity," m, which is equal to s/w, Table I.

The slope of the curve itself at the same point is the "tangent relative sensitivity," n, which is equal to $\partial s/\partial w$. Both m and n approach the initial relative sensitivity value, M, as the relative displacement approaches zero, this property of these functions provides a convenient definition for two useful alinearity coefficients.

C. Gauge Alinearity

Two alinearity coefficients, a and b, can be derived from the two displacement sensitivity factors, m and n. The "secant alinearity coefficient," a, is defined as the ratio of the secant sensitivity to the initial sensitivity, m/M, and the "tangent alinearity coefficient," b, as n/M. Both of these coefficients are defined for a constant value of the radius ratio, B/A, and are a function of w.

These two coefficients have the value of 1.00 as the relative displacement of the diaphragm approaches zero, i.e., a capacitor gauge is most linear for those diaphragm deflection values which are a very small fraction of the air gap value. This relation is true for the opposite cases of recession and approach, for both types of diaphragm; in recession, the coefficient values are less than 1.00; in approach, greater than 1.00. Figures 8 through 13 show these relations between a, b, s, w, and B/A for the two diaphragms.

If the displacement of the center of the diaphragm, y_0, is a linear function of the volume change, v, or of the applied pressure, p, both coefficients give the alinearity values for actual

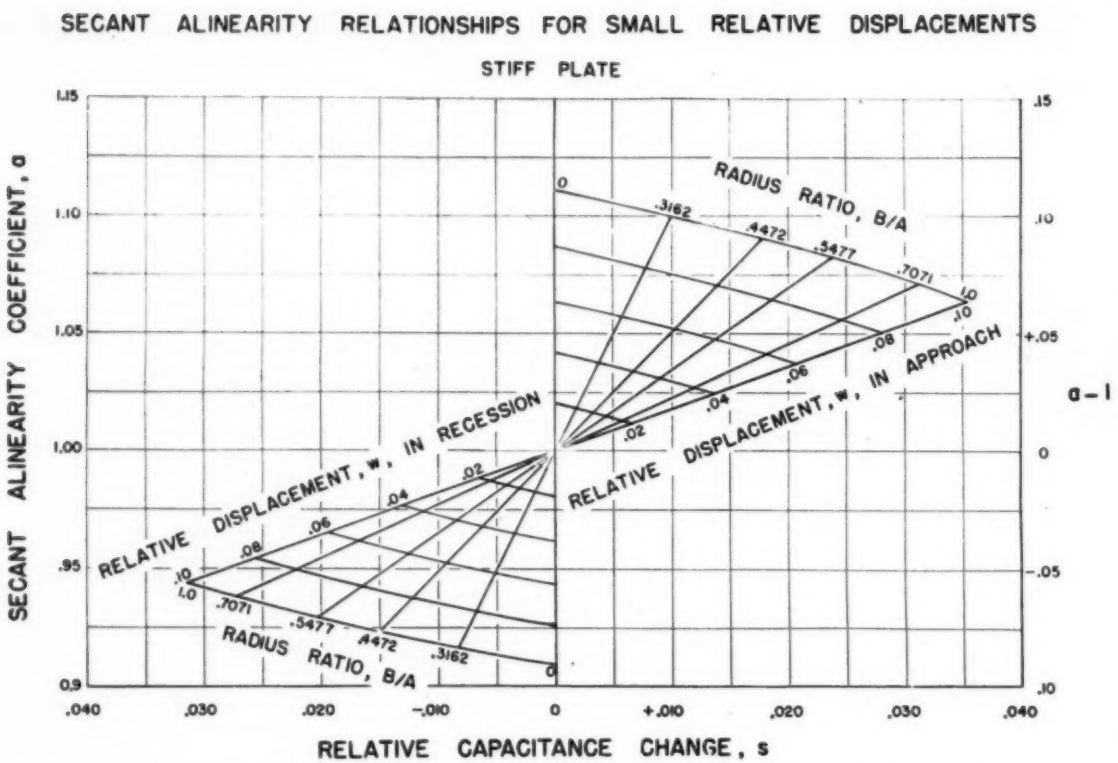


FIG. 9. The secant alinearity coefficient relationships for a stiff plate for small values of the relative displacement. This figure is an amplification of the part of Fig. 8 near zero capacitance change. In those uses of the gauge requiring a high degree of linearity, the variables are restricted to the range on this plot or to an even smaller range. The considerations mentioned under Fig. 8 apply also to this figure.

volume and pressure gauges. In other words, for a constructed gauge the ratio of the slope of the volume change-capacitance curve, or of the pressure-capacitance curve at a given pressure or volume, to the initial slope is given directly by the value of the alinearity coefficients at a given s, the relative capacity signal for a given radius ratio, B/A. A mechanical design which has a fairly linear pressure-deflection relation is discussed in the section on experimental performance.

D. Absolute Types of Sensitivity

The absolute sensitivity factors are divided into the two general types, secant and tangent, defined above for the relative displacement factors.

It is convenient to define each of the secant factors as the ratio of the capacitance signal, c, to the displacement, y_0, to the volume, v, or to the pressure, p. The tangent sensitivities are defined as the rate of change of the capacitance with respect to each of the three variables. All of these absolute factors are calculable from the gauge constants, A, B, D (or C_0) and the corresponding relative displacement sensitivities, m and n; however, it may be easier to use these constants, the initial relative displacement sensitivity, M, and the two alinearity coefficients, a and b. The equations for these calculations are given in Table II.

The absolute displacement sensitivity, (c/y_0), for a given value of B/A, can be shown to be equal to one-fourth the product of the square of the ratio of the diaphragm radius to the air gap, $(A/D)^2$, and the relative sensitivity, M (Table II); c/y_0 has an initial value of $(c/y_0)_0$ when m equals M. The corresponding tangent factor, dc/dy_0, is similarly calculable by substitution of n for m (Table II). The values of a for these and later calculations for various values of B/A and w are given in Figs. 8, 9, and 10; the b values, in Figs. 11, 12, and 13.

In other words, the absolute displacement

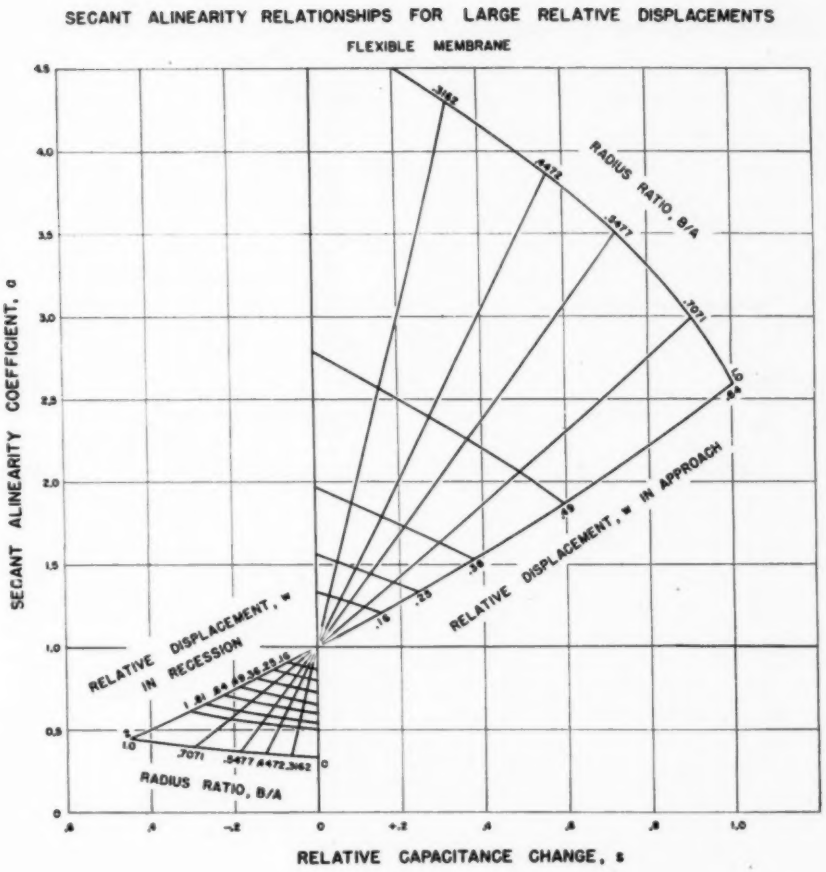


FIG. 10. The secant alinearity relationships for a flexible membrane. This plot for a flexible membrane is of the same type as the one for the stiff plate shown in Fig. 8. Comparing this figure with number 8 it can be seen that the flexible membrane has greater deviations from ideal linearity (a equal to 1.0) than does the stiff plate for the same value of the relative displacement, w.

sensitivities, secant and tangent, are independent of the absolute values of the diaphragm radius, thickness, and material; of the electrode radius; and of the air gap. However, these factors are dependent on the ratio of the diaphragm radius to the air gap, and on the relative displacement sensitivity.

The absolute volume sensitivity factors, for a given value of B/A are calculable by the equations given in item 22, Table II; or as in the case of c/y_0, by the substitution of M and a or b, for m or n in the equations. By these equations it can be shown that the two absolute volume sensitivities for both diaphragm types are independent of the absolute value of the diaphragm radius, thickness, and material; and of the electrode radius. However, the volume sensitivity varies inversely with the absolute value of the initial air gap. This result suggests that the air gap be designed and constructed with as small a value as is practicable in those gauge applications in which the volume change is desired minimal.

There are two alternative ways of calculating the absolute pressure sensitivities; the equations are given as items 23 and 23a in Table II. The first method (item 23) in the stiff-plate case involves knowing at least two of the elastic constants of the material of the diaphragm (Table I, item 14; Young's modulus and Poisson's ratio) with sufficient accuracy for the design problem, and using these to calculate the deflection constant.[11] In the case of the flexible membrane, it involves knowing the value of the initial stretch tension ($G/4$, item 15, Table I) in the membrane to calculate the deflection constant.[9] It was quickly found that the literature lacks the data necessary for a highly accurate calculation of the deflection constant for a stiff plate; and that the methods of measuring or calculating the initial stretch were unsatisfactory in the case of the flexible membrane. However,

[11] R. J. Roark, *Formulas for Stress and Strain* (McGraw-Hill Book Company, Inc., New York, 1938), p. 171, Case 6.

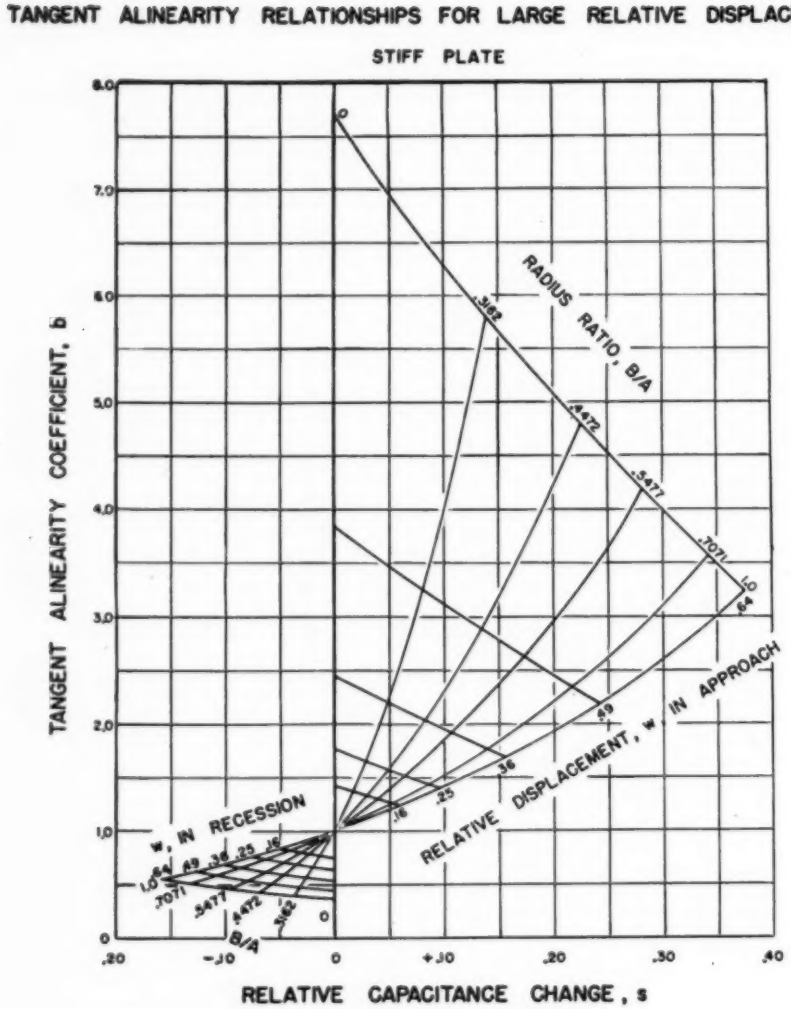


FIG. 11. The tangent alinearity coefficient for a stiff plate for large relative displacements. The tangent alinearity coefficient, b, (Table II), is defined as the ratio of the tangent relative displacement sensitivity, n, to the initial sensitivity M. A comparison of this plot with Fig. 8 shows the large deviations of the variational sensitivity (n) from linearity compared with the secant sensitivity, m.

for a first approximation in a design problem, this first method is useful.

The second method (item 23a, Table II) of calculating the pressure sensitivity envolves an independent measurement of y_0/p with a diaphragm of the desired material, thickness, and radius. Figure 14 shows such a measurement on a stainless steel diaphragm by an interferometer method. With the determined value of y_0/p, the deflection constant can be calculated (Table I) and used in further design modifications to obtain any desired c/p in an actual gauge. The data of Fig. 14 were used to calculate theoretical values of c/p in the evaluation of this design theory (Fig. 15).

EVALUATION OF THE DESIGN THEORY

Using a stiff-plate gauge shown in Figs. 16 through 19, data were taken on the pressure-capacitance performance (c/p ratios) in the approach and recession conditions (Fig. 15). To check these data against the theory, theoretical c/p values were calculated from the equation in item 23a (Table II) using the measured values of the electrode radius, B, the diaphragm radius, A, the initial capacitance, C_0, the displacement-pressure ratio, y_0/p, the value of M from Fig. 6, and the values of a from Figs. 8 and 9. The values of the y_0/p ratio were determined by an interferometric method (Fig. 14), rather than by the less accurate method of calculation from the elastic constants. The results of these calculations are plotted on the graph of Fig. 15 with the experimental data. Figure 15 shows that the stiff-plate theory can be used with a fair degree of accuracy for design purposes, even in the

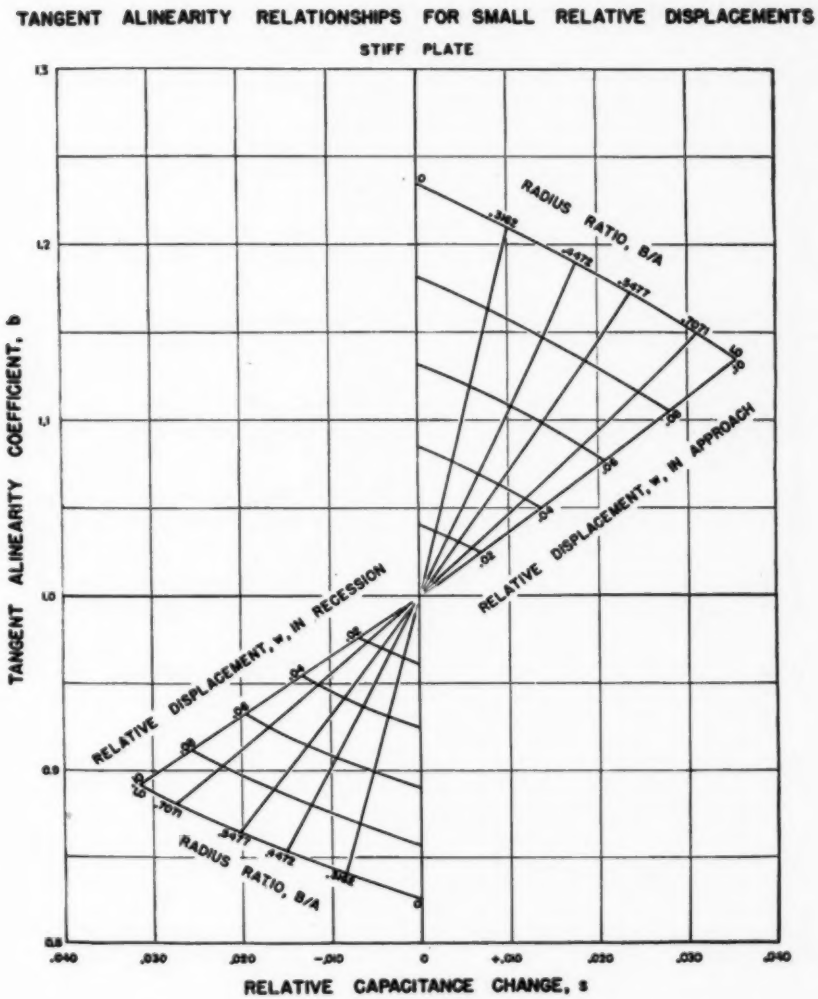

TANGENT ALINEARITY RELATIONSHIPS FOR SMALL RELATIVE DISPLACEMENTS
STIFF PLATE
TANGENT ALINEARITY COEFFICIENT, b
RELATIVE CAPACITANCE CHANGE, s
RADIUS RATIO, B/A
RELATIVE DISPLACEMENT, w, IN APPROACH
RELATIVE DISPLACEMENT, w, IN RECESSION

FIG. 12. The tangent alinearity coefficient for a stiff plate for small relative displacements. This plot is an enlarged view of Fig. 11 near the relative displacement of zero magnitude. This is a companion figure to number 9 for designs involving a high degree of linearity.

extreme cases involving large relative displacements ($w = \pm 0.49$) in approach and in recession. The flexible membrane theory still awaits confirmation; it is hoped that this can be done and reported in the near future; there are a few unpublished data which show that reasonably accurate predictions of performance can be expected from the flexible membrane theory.

SPEED OF RESPONSE OF A CONSTRUCTED GAUGE

Figure 20 shows the response of a gauge with an attached needle to a sudden drop in pressure level. Details of the technique of obtaining the records is given under the figure; the gauge itself is pictured in Figs. 17, 18, and 19. These records show the importance of having a homogeneous conducting medium between the source and the gauge diaphragm. They also show that fairly high speed pressure recording can be accomplished through small conduction tubing with a relatively incompressible fluid, if the volume change of the gauge is small and if the tubing wall is relatively stiff. The volume change (Table 1) for the full pressure range in Fig. 20 is approximately 2×10^{-5} cc. The contained liquid volume is 0.4 cc; therefore an additional 20 percent must be added to the gauge volume change to allow for the compression of the contained water at a pressure of 150-mm Hg above 1 atmosphere.

For the speed of response of the gauge without an attached fluid system it was found that it was difficult to obtain a step change in pressure level sufficiently fast to approach the lowest natural mode of vibration of the stiff-plate diaphragms used (10 to 60 kilocycles). However, the applica-

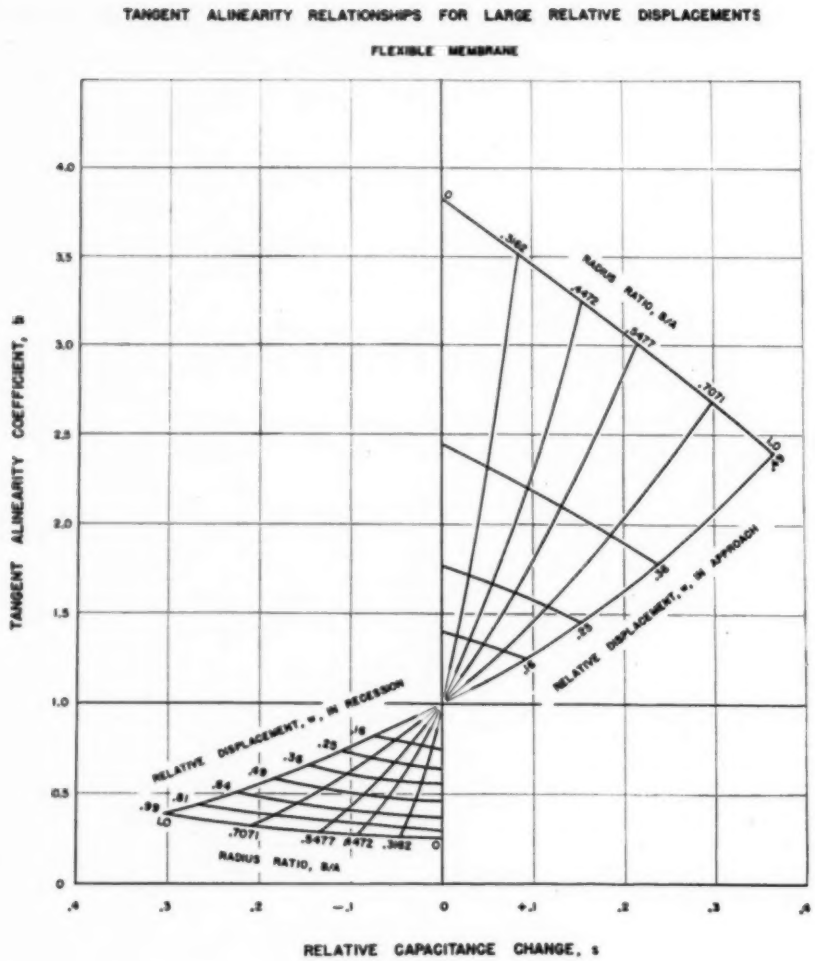

TANGENT ALINEARITY RELATIONSHIPS FOR LARGE RELATIVE DISPLACEMENTS
FLEXIBLE MEMBRANE
TANGENT ALINEARITY COEFFICIENT, b
RELATIVE CAPACITANCE CHANGE, s
RADIUS RATIO, B/A
RELATIVE DISPLACEMENT, w, IN APPROACH
RELATIVE DISPLACEMENT, w, IN RECESSION

FIG. 13. The tangent alinearity coefficient for a flexible membrane. This figure is to be compared to Fig. 11 for a stiff plate. The tangent alinearity coefficients of the two cases do not differ to a great degree, in contrast to the secant alinearity coefficients (Figs. 8 and 10).

tions of the gauge to date have not made full use of the available speed of response.

CONSTRUCTION DETAILS TO OBTAIN SMALL AIR GAPS

Standard machining is performed on all components (Fig. 16) except those having critical surfaces which require special techniques as follows:

In the electrode barrel assembly, the electrode (2, Fig. 16) and its lava insulators (6) are assembled to the barrel (4) with bakelite laquer, riveted (12, Fig. 17) and baked. This assembly is faced in the lathe and hand-lapped on a plate-glass lap, until the face of both the electrode and its barrel are in the same flat plane. The electrode spacing is generated by hand-lapping the electrode face with a brass lap. The amount of spacing thus generated is gauged by placing the assembly against a flat metal gauge block and measuring the electrical capacitance between the electrode and the gauge block. If the desired spacing is exceeded, correction may be made by lapping the face of the barrel on the plate glass.

For a given diameter, diaphragms are handled differently as they vary in thickness. A diaphragm of $\frac{11}{16}''$ outside diameter could be classified thick if it were 0.008" and over, medium 0.004" to 0.008", thin 0.004" and under.

Thick diaphragms are machined from bar stock and hand-lapped to size on plate glass. Frequent turning of the diaphragm, exposing first one face then the other to the lap, will minimize warping from strain release. The thickness of the diaphragm is gauged with a micrometer. The flatness is checked by optical means or by substituting the diaphragm for the gauge block in the capacitance test described above.

Medium diaphragms are disks cut from polished, hard rolled strip metal of the required thickness. The disks are generally curved in a cylindrical plane but approach a flat surface when clamped in position.

Thin diaphragms are disks cut from polished, half-hard rolled metal of the desired thickness. They require stretching while clamping to present flat surfaces. The stretching mechanism (not shown) consists of a recess in the electrode barrel face into which the diaphragm is pressed by a ring member. The ring member is designed with clearance to insure positive seating of the diaphragm.

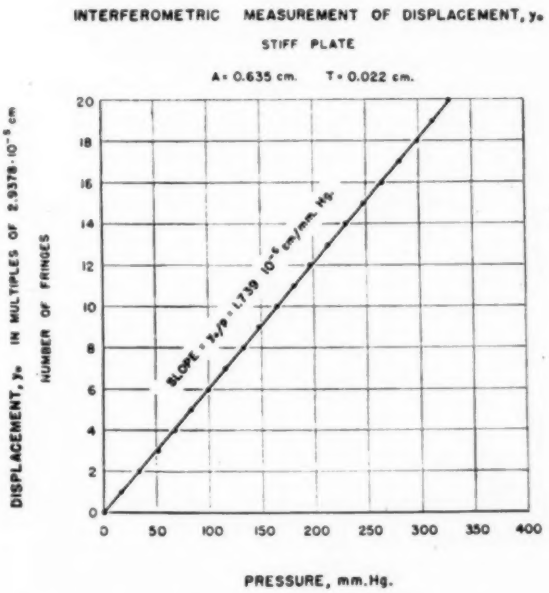

INTERFEROMETRIC MEASUREMENT OF DISPLACEMENT, y_0
STIFF PLATE
A = 0.635 cm. T = 0.022 cm.
DISPLACEMENT, y_0 IN MULTIPLES OF $2.9378 \cdot 10^{-5}$ cm
NUMBER OF FRINGES
SLOPE = $y_0/p = 1.739 \cdot 10^{-4}$ cm/mm. Hg.
PRESSURE, mm.Hg.

FIG. 14. Experimental determination of the pressure-displacement relation for a stiff plate. The plate is edge-clamped, and corresponds to the diaphragm of the type shown in Fig. 2. The capacitance-pressure relations for this unit are presented in Fig. 15, and the unit itself in Figs. 17, 18, 19, and 20. The absolute central diaphragm displacement, y_0 (Table I) was measured with a vertical incidence Newton's fringe type of interferometer, using a yellow line (5875.6 Angstroms) of a helium discharge tube. The fully aluminized plate (1×3×3 mm) of the interferometer was moved toward the 60 percent aluminized fixed plate by displacement of the diaphragm's center by an applied pressure. The pressure was measured with a compensated mercury manometer with an accuracy of ±0.1 mm Hg. The fringes were counted by means of a microscope and an eyepiece scale. Each fringe counted represents a movement of the diaphragm center of one-half of a wave-length of the He line, given on the left of the figure. The pressure range usually used with this gauge unit is that represented by the first two fringes, or 30 mm Hg. The diaphragm displacement is away from the electrode, i.e., in the recession condition.

The strict proportionality between the displacement and the applied pressure up to at least 300 mm Hg (ten times the range actually used) shows that this diaphragm is operated well inside of the proportional elastic limit.

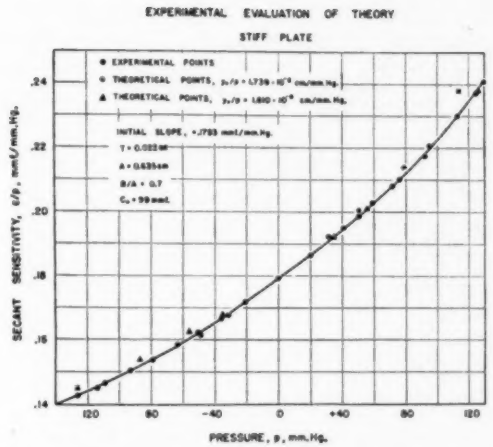

EXPERIMENTAL EVALUATION OF THEORY
STIFF PLATE
EXPERIMENTAL POINTS
THEORETICAL POINTS, $y_0/p = 1.739 \cdot 10^{-4}$ cm/mm. Hg.
THEORETICAL POINTS, $y_0/p = 1.810 \cdot 10^{-4}$ cm/mm. Hg.
INITIAL SLOPE, .1793 mmf/mm.Hg.
T = 0.022 cm
A = 0.635 cm
B/A = 0.7
C_0 = 99 mmf.
SECANT SENSITIVITY, c/p, mmf/mm.Hg.
PRESSURE, p, mm.Hg.

FIG. 15. An experimental evaluation of the design theory for a stiff plate. The capacitance change, c, and the initial capacitance, C_0, were measured with a "Q" meter at a frequency of 3.0 megacycles and by substitution of a precision condenser (General Radio Company). The pressure was measured with the Hg manometer mentioned under Fig. 14. The absolute pressure secant sensitivity factor, c/p, was calculated from the measured values of c and p; the solid dots and the drawn curve represent these experimental data. The dotted circles are theoretical points calculated from the value of the recession state y_0/p given in Fig. 14, the B/A value of 0.7 (measured B and A), the C_0 value of 99 mmf, the M value of Fig. 6, and the secant alinearity coefficient values of Fig. 8, by means of the equation given as item 23a in Table II. It can be seen that in recession these points fit the experimental curve fairly well, but deviate from the curve in the high pressure region in the approach state. It was then found, by the interferometric method of Fig. 14, that the value of y_0/p in the approach direction was smaller than in recession (value given in this figure). The points calculated with this value of y_0/p are the triangles. Presumably the difference of y_0/p in the two directions is due to unsymmetrical clamping supports on the two sides of the plate.

The main body (3, Fig. 16) has an interior seat against which the diaphragm is clamped. This seat is hand-lapped to a flat surface with a formed brass lap. The exterior faces of the main body are hand-lapped on plate glass to present flat seats to the mating attachments (13 and 15, Fig. 17), which are similarly finished.

All lapping operations referred to above were performed with a fine grain abrasive and light oil. Silicon carbide No. 400B was used for coarse grinding and No. 600B for finish.

ELECTRICAL CIRCUITS FOR USE WITH THE GAUGE

The circuit actually used with these gauges will be published later. The choice of a given

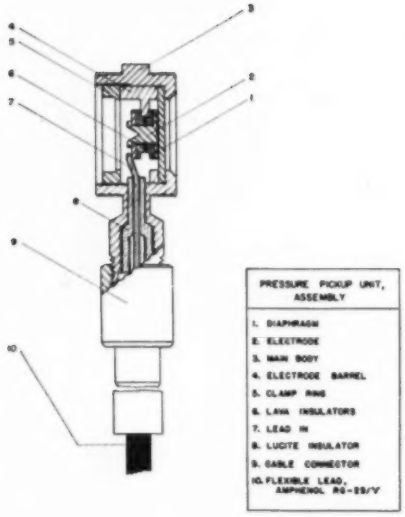


FIG. 16. Sectional drawing of a variable capacitor. The active diaphragm (1) diameter is 0.50 inches. The radius ratio B/A is 0.70. The electrode-diaphragm air gap is 5.10^{-4} cm. (0.0002 inch). The pertinent parts are labeled in the figure. The lead-in (7) carries the r.f. voltage to the electrode; the rest of this unit is grounded. It is to be noted that the electrode side of the diaphragm is kept as open as possible to allow quick venting of the gas in this space in applications in which large ambient pressure changes occur, such as in the explosive decompression of pressurized aircraft cabins in altitude chambers. The lava insulators (6) are supported by three equally spaced legs protruding inward radially from the electrode barrel (5); only one of these legs shows in this section. This construction allows the quick venting mentioned above. The construction details to assure small air gaps are given in the text.

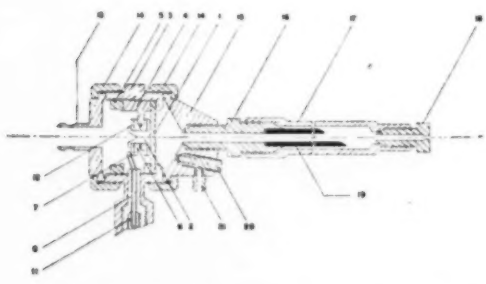


FIG. 17. Sectional drawing of a capacitor with an attached liquid system. Through number 8, the parts are the corresponding ones of Fig. 17. Photographs of this unit are shown in Figs. 18 and 19. A tubing attachment (13) for pressure calibration purposes is shown fastened to the electrode side of the unit by the retaining ring (14). A liquid system (15) with an attached hypodermic needle (16 and 19) is fastened to the diaphragm side of the unit. For the sake of clear reproduction the needle is greatly enlarged; the true relative size is shown in Fig. 18; a side tube (21) and a needle valve (20) are used for filling the liquid system. A needle guard (17) and its terminal plug (18) keep the needle wet and sterile until ready for use. This type of unit has been used for blood pressure, intrathoracic pressure and intraspinal pressure recording (Figs. 21 through 24); its response is shown in Fig. 20.

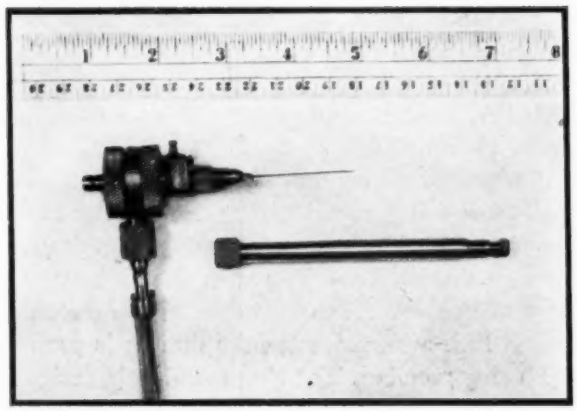


FIG. 18. Internal body pressure condenser gauge. A cross section drawing of this unit and its details is shown in Fig. 17. A photograph of this unit and its attachments is given in Fig. 19. The needle guard is below the needle. This unit is used for blood pressure and other internal body pressure recording (Figs. 23 and 24).

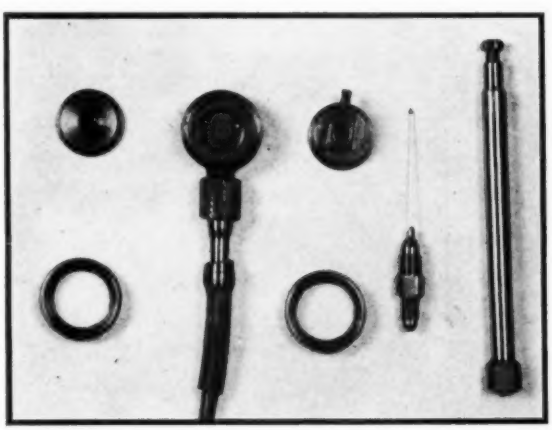


FIG. 19. Pressure gauge with attachments. This is the gauge of Fig. 18 with the tubing connector (left), the liquid system, the needle and its guard (right), removed. The one-half-inch diameter active diaphragm surface is seen as a central ring on the face of the pickup unit. All parts and the gauge itself are made of stainless steel.

circuit depends on the application; several satisfactory ones have been published.[8, 9, 12, 13] One requirement for biological work is that the electrical system record sustained capacitance changes, i.e., record the direct current component as well as the alternating current ones. Cathode-ray oscillographs and oscillograph galvanometers have both been used with success for recorders.

[12] D. W. Dana, Rev. Sci. Inst. **5**, 38 (1934).
[13] C. H. Brookes-Smith, J. Sci. Inst. **16**, 361 (1939).

APPLICATIONS OF THE GAUGE TO MECHANICAL MEASUREMENTS IN BIOLOGICAL SYSTEMS

Various applications of the gauge to pressure recording during the war are presented in Figs. 21 through 25. Since the applications are described under the figures, no further comments will be made here. Other uses of the gauge as an ultramicrometer or a dilatometer in biological systems might be mentioned. The gauge may be used to measure the movements of the body resulting from recoil caused by heart and blood motions in an electrical ballistocardiograph. An obvious use is as a high speed, isometric "lever" for recording the tension developed by muscles during fast or sustained contractions; it may be possible to use it for single muscle fibers. As a dilatometer, the gauge can be used in plesthysmography, thermometry, and respirometry. The volume changes in muscle during fast contractions may possibly be recorded. It may be used in the measurement

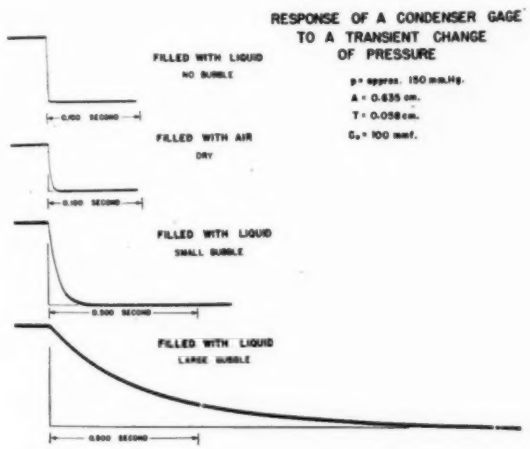


FIG. 20. Response of a capacitor gauge with attached needle to a sudden fall in pressure level. The gauge itself is of the type shown in Figs. 18 and 19. The attached hypodermic needle was a number 23, one-inch long. The volume of the liquid system is 0.4 cc. The needle was immersed in liquid or in air in a short $\frac{1}{2}$-inch diameter well in the face of a 6-inch diameter, horizontal plate, over which was stretched a thin paper diaphragm. The pressure in the air space between the plate and the paper diaphragm was raised with an air supply, and the paper diaphragm was "exploded" away with a mouse trap lever traveling at high speed. The paper broke completely away over the whole plate surface, causing a rapid fall in pressure over the well containing the needle. The conditions causing the various rates of fall of the gauge response are given in the figure. The top record shows a response time of about one-thousandth of a second, which is an upper limit, presumably set by a time constant in the electrical circuit used for recording, rather than by the response time of the mechanical system. (To fill the system with liquid without bubbles, the contained air was flushed out with pure carbon dioxide gas before filling; any bubbles formed of CO_2 quickly dissolve in the liquid, leaving the system completely liquid filled.)

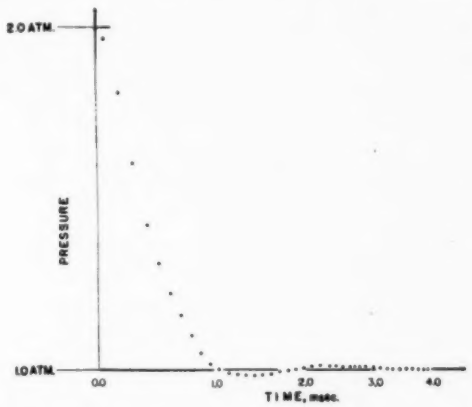


FIG. 21. Explosive decompression of a short cavity. The condenser pressure gauge (of the type shown in Fig. 17) was placed in the bottom of a well 3 inches long and $\frac{3}{4}$-inch in diameter and containing air. By the technique described under Fig. 20, the pressure at the mouth of this well was suddenly dropped to the ambient pressure from twice the ambient value. This retouched record of the pressure was photographed on a cathode-ray oscilloscope. The time between dots is one ten-thousandth of a second. Note the relatively small after-vibrations in contrast with those of Fig. 22. This gauge was subsequently used in recording the pressure fall in explosively decompressed pressurized cabins in an altitude chamber; the minimum time of pressure fall recorded in these cabins was about four times the value shown here.

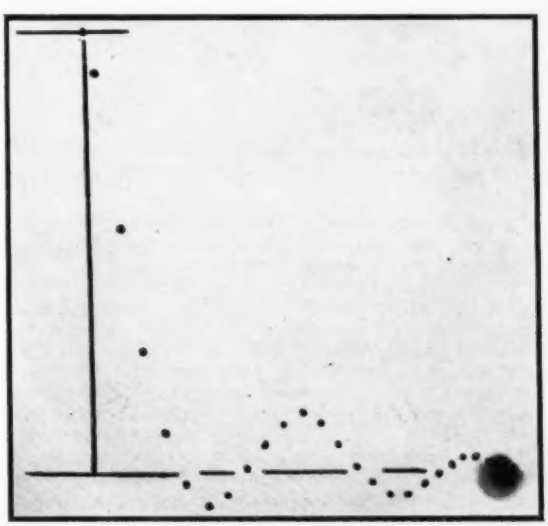


FIG. 22. Explosive decompression of a long air cavity. The technique and the pressure range are the same as those of Fig. 21. The $\frac{3}{4}$ inch diameter cavity in this case, however, is 3 feet long. The time between dots is two ten-thousandths of a second. The amplitude of the first after-vibration is 2.0 pounds per square inch, after an initial fall of 15 pounds per square inch. This record shows some of the artifacts created by trying to measure relatively large rapid changes in pressure through a long pressure lead, filled with air.

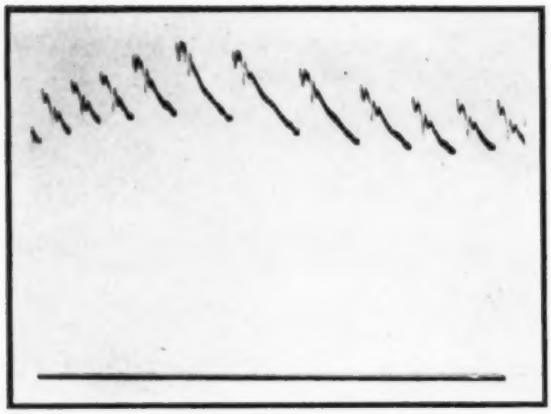


FIG. 23. A blood pressure record in the carotid artery of a dog. The pressure base line is at the bottom of the figure. The peak pressure is approximately 150 mm Hg. The record was taken with a pressure gauge of the type shown in Figs. 18 and 19. This record is shown through the kindness of H. C. Bazett.

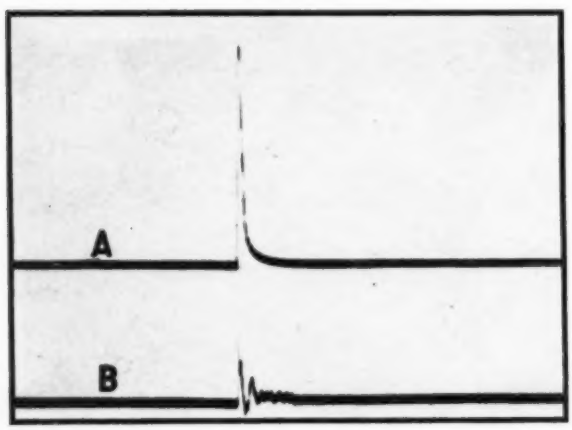


FIG. 24. Record of the intrathoracic and the intraspinal pressure in an explosively decompressed dog. *A* is the intrathoracic pressure record, and *B* the intraspinal pressure record. The breaks in the falling phase of record *A* are at 1/120 second intervals. The condenser gauge used was of the type shown in Figs. 18 and 19. The dog was in a small chamber at an equivalent pressure of 8000 feet; a thin film diaphragm separated this chamber from a larger one at a lower pressure. When the diaphragm was ruptured, the air pressure around the dog fell in 0.03 seconds to that equivalent to 50,000 feet. The two gauge pick-up units were liquid filled, with an electrode side open to ambient pressure and its changes (Fig. 16). The needles were in their respective body spaces. The pressure rise with respect to ambient (about 30 mm Hg in *A*) is apparently due to the relative delay of the air flowing from the lungs. This figure is reproduced through the courtesy of Fred A. Hitchcock, Ohio State University.

of the stimulus of nerve or muscle when pressure or mechanical stresses are applied in a slow or

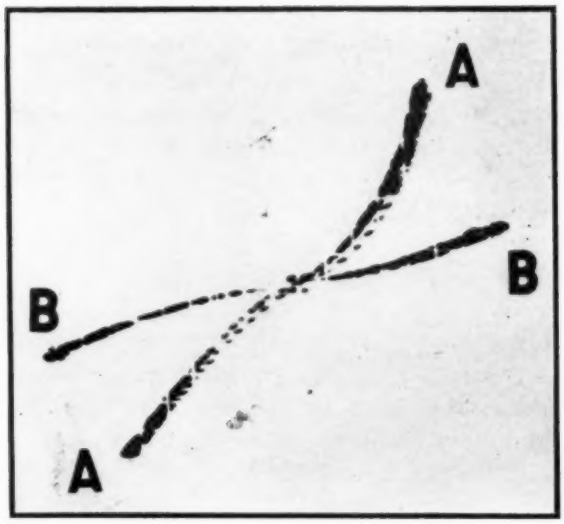


FIG. 25. Pressure-flow curves for air through the nasal passageway. Two, flexible membrane-type pressure gauges with a pressure range of 100 mm of water and 10 mm of water for a capacitance change of 5 mmf were used for recording air pressure and air flow, respectively. The record is a photograph of a cathode-ray tube screen, with flow appearing on the horizontal axis and pressure on the vertical axis. Zero flow and zero pressure are at the center of the record where the two lines cross. Expired gas flow and pressure are in the right upper quadrant, inspired values in the lower left one. The flow values cover about ±50 liters/minute; the pressure about ±80 mm of water. The flow is recorded by measuring the pressure drop across a laminar-flow type glass wool resistor, placed in the tube exit from an oxygen mask covering the subject's nose and mouth. The pressure was measured differentially between the inside of the mask and in *A–A*, the subject's oropharynx, and in *B–B*, his nasopharynx. In both cases, one nostril is plugged; in *A–A* the pressure tap tube goes into the mouth; in *B–B* the tube goes through the plugged nostril. Record courtesy of Forman and Benevides.

in a fast manner. It is hoped that the design theory presented above will be an aid in extending the applications of this type of gauge in biological research and in the engineering and physical fields.

ACKNOWLEDGMENTS

The authors wish to express their thanks to Dr. D. W. Bronk for his encouragement during the course of this work; to A. J. Rawson and J. P. Hervey for their help and suggestions; and to J. R. Pappenheimer and G. A. Millikan for their stimulating applications for this gauge.

Conductivity of Metallic Surfaces at Microwave Frequencies*

E. MAXWELL
Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received February 4, 1947)

Methods of measuring effective conductivities at microwave frequencies are described. These consist of either measuring the transmission loss in a long waveguide, or in measuring the Q's of resonant cavities. Both methods have been applied to measurements at 1.25 cm. Results for a number of metals are presented. Deviations from d.c. conductivity are thought to be due to surface roughness.

1. INTRODUCTION

IN the course of development of microwave components it was early observed that the high frequency conductivity of many metallic surfaces was frequently lower than the d.c. conductivity. This phenomenon was especially noticeable in the 1.25-cm region. The work described here was done for the purpose of collecting data on the conductivities of representative surfaces in the region of 1.25 cm.

The r-f conductivity is important in at least two applications. These are:

(1) Long transmission systems where the total attenuation is governed by the conductivity.

(2) Cavity resonators, in which the maximum attainable Q is limited by the effective conductivity.

The effective conductivity of metals at microwave frequencies is governed by a superficially thin surface layer. The current density inside the metal decreases exponentially with depth according to $J=J_0\exp-[(\omega\mu\sigma/2)]^{\frac{1}{2}}x$, where J_0=current density at the surface, ω=angular frequency, μ=permeability, σ=conductivity, and x=distance into conductor measured along the normal.

When $x=(2/\omega\mu\sigma)^{\frac{1}{2}}$, $J=J_0/e$. This value of x is defined as the skin depth δ. It is easy to show that the total loss, obtained by integrating $J^2/2\sigma$ with respect to x, (between 0 and ∞), is equal to $J_0^2/2\sigma\delta$. For purposes of calculating losses one may consider the current confined to a skin of thickness δ within which the current density is uniform and equal to J_0. Actually of course 37 percent of the total current flows at depths below δ.

This formulation is rigorously correct only for an infinite plane surface. However, it remains an excellent approximation for curved surfaces whose radii of curvature are large compared to δ. In computing losses one evaluates the integral

$$\int\int_S \frac{J_0^2}{2\sigma}ds=\int\int_S \frac{H_t^2}{2\sigma}ds$$

taken over the surface of the waveguide or cavity. H_t is the tangential component of magnetic field. (The factor 2 comes in since we are interested in average power dissipation, and J_0 and H_t are both maximum values of sinusoidal time functions.) If the metal surface is not smooth but has surface irregularities and scratches whose dimensions are comparable with δ, the foregoing treatment is evidently an over-simplification. The presence of scratches may alter the current distribution especially if the scratches are in a direction orthogonal to the lines of current flow. The scratches could conceivably act like constrictions in the effective cross section, thereby increasing the resistance.

2. MEASURING TECHNIQUES

Two general methods of measurement have been employed. In the first method the sample to be measured is in the form of a long piece of waveguide or equivalent transmission system. If one end of the waveguide is shorted, the standing-wave ratio seen looking into the other end is a known function of the guide attenuation, which in turn depends on the conductivity. Thus, from a single measurement of standing-wave

* The research reported in this paper was made possible in part through support extended the Massachusetts Institute of Technology, Research Laboratory of Electronics, jointly by the Army Signal Corps, the Navy Department (Office of Naval Research), and the Army Air Forces (Air Materiel Command), under the Signal Corps Contract No. W-36-039 sc-32037.

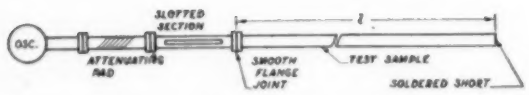


FIG. 1. Attenuation measurement by long guide technique. Rectangular guide.

ratio, the effective conductivity may be computed.

In the second method, the sample forms the walls of a resonant cavity. The Q of the cavity is a known function of the conductivity, and hence the conductivity may be calculated from the measured Q. A variation of this method was found quite useful. The input conductance at resonance is directly proportional to the losses in the cavity, or more precisely, to the square root of the wall resistivity. The constant of proportionality depends on the coupling and, if known, a single measurement of the input standing-wave ratio at resonance permits one to calculate the resistivity.

2.1 DETAILS OF FIRST METHOD

The measurement set-up for rectangular wave guide is illustrated in Fig. 1. If r is the measured voltage standing-wave ratio, and l is the length of the wave guide, the attenuation is given as follows:

$$\text{Attenuation per unit length} = \alpha$$

$$= -\frac{10}{l}\log_{10}\left(\frac{r-1}{r+1}\right) \text{ db}$$

$$= -\frac{1}{2l}\log_e\left(\frac{r-1}{r+1}\right) \text{ nepers.}$$

The attenuation and conductivity are related by the following expression:

$$\alpha = \frac{1}{2b}\cdot\frac{1}{[1-(\lambda_0/\lambda_c)^2]^{\frac{1}{2}}}\cdot\left(\frac{4\pi}{\lambda_0\mu c\sigma}\right)^{\frac{1}{2}}\cdot\left[1+\frac{2b}{a}\left(\frac{\lambda_0}{\lambda_c}\right)^2\right],$$

where α = attenuation in nepers per meter, σ = conductivity in mhos per meter, a = wave guide width in meters, b = waveguide height in meters, λ_0 = free space wave-length in meters, $\lambda_c = 2a$ = guide cut-off wave-length in meters, μ_0 = permeability of free space (M.K.S.) $= 4\pi \times 10^{-7}$, and c = velocity of light $= 3\times 10^8$ meters per second. This expression is for the TE_{10} mode only. For $0.170'' \times 0.420''$ wave guide at 1.25 cm, the formula reduces to:

$$\alpha = 299/(\sigma)^{\frac{1}{2}} \text{ nepers per meter.}$$

The length of waveguide is conveniently chosen so as to have a voltage standing-wave ratio of about 5 or 6. At 1.25 cm with common materials, this usually means a waveguide length of the order of six feet. The correction for losses in the slotted section is then negligible.

The far end of the waveguide should be shorted by means of a soldered plate. A choke plunger is unsatisfactory for this purpose, since some power may leak past. If the waveguide is suspected of inhomogeneities or bad spots, a quarter wave-length should be removed from the shorted end and the test repeated. This will shift the standing-wave pattern by a quarter wave-length, and if bad spots are present, the measured loss will probably be different in the two cases. The loss in the short circuiting plate is of course negligible in comparison to that in the wave guide itself.

The foregoing procedure is straightforward enough when dealing with rectangular wave guide. Frequently, however, one wishes to use circular waveguide. This can be done by using a round-to-rectangular transition piece, or transformer, between the round pipe and the slotted section, but a special technique must be employed to avoid errors due to elliptical polarization in the pipe.

It is inevitable that any long piece of round pipe will have both ellipticity and skewness which will cause the incident wave to split up into a pair of cross-polarized waves. Each of these cross-polarized waves will be reflected at the short and come back down the pipe, but only the vertical component of each will couple into the rectangular waveguide. This will cause the observed standing-wave ratio in the rectangular slotted section to be less than the true standing-wave ratio at the sending end of the round pipe.

This difficulty may be eliminated as follows. It may be shown theoretically that given any skewed elliptical pipe, it is possible to choose a direction of polarization for the incident wave, such that the reflected wave arrives with that same polarization. The reflected wave then couples completely into the rectangular wave

guide. and the standing-wave ratio observed in the slotted section is the true standing-wave ratio.

If the pipe is a perfect elliptic cylinder, it is clear that either of the principal axes of the ellipse may be taken as the preferred direction. An actual pipe is more apt to resemble a twisted cylinder in which both the orientation of the axes and the ellipticity are functions of position along the axis of propagation. Thus there may be cross-coupling between the modes, and the reflected wave will not generally have the same orientation as the incident wave.

In practice the round wave guide is arranged so that it may be rotated with respect to the transformer (see Fig. 2). The wave guide is rotated until the position of the maximum standing-wave ratio is found. This is the condition for which the reflected wave returns with the same polarization as the incident wave. A convenient type of round-to-rectangular transformer is a quarter wave-length of oval cross section intermediate in shape between the round and rectangular sections. The loss in such a transformer is quite negligible.

This technique was applied in measuring the effective conductivity of mercury. A circular waveguide was constructed by immersing a polystyrene rod in mercury. The details of this experiment are described in the Appendix.

2.2 DETAILS OF SECOND METHOD

In any resonant cavity, the unloaded Q is given by

$$Q_0 = \text{const.} \times \frac{\lambda}{\delta},$$

where λ is the resonant wave-length, and δ is the skin depth. The constant of proportionality is a function of the resonator geometry and is known for many of the simpler cases. Thus if one measures Q_0, it is simple to compute the effective conductivity from δ (if the resonator geometry is one for which the theoretical problem has been solved).

In practice, one couples a waveguide into the cavity by means of a coupling window of some sort and measures the variation of input standing-wave ratio with frequency. The half width

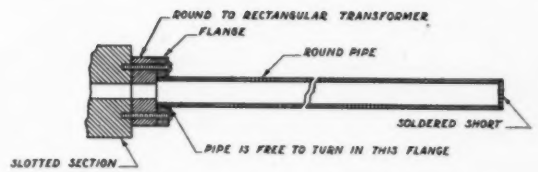


FIG. 2. Attenuation measurement by long guide technique. Round guide.

of the resulting resonance curve determines Q_L, the loaded Q. Thus,

$$\frac{1}{2Q_L} = \frac{\omega_{\frac{1}{2}} - \omega_0}{\omega_0} \tag{1}$$

where ω_0 is the resonant frequency and $\omega_{\frac{1}{2}}$ the frequency corresponding to the half-power point. (The half-power point is the frequency at which the power absorbed by the cavity is half the power absorbed at resonance.) The ordinates of the half-power points are given by

$$\beta_{\frac{1}{2}} = \frac{\beta_0 + 1 + (\beta_0^2 + 1)^{\frac{1}{2}}}{\beta_0 + 1 - (\beta_0^2 + 1)^{\frac{1}{2}}}, \tag{2}$$

where β_0 = the voltage standing wave ratio at resonance, and $\beta_{\frac{1}{2}}$ = the voltage standing wave ratio at the half-power point.

The loaded and unloaded Q's are related by

$$\frac{Q_0}{Q_L} = 1 + \frac{Y_0}{G}, \tag{3}$$

where Y_0 = the line admittance, and G = the input conductance of the cavity. If the cavity is undercoupled,

$$Y_0/G = 1/\beta_0 \tag{4a}$$

and if overcoupled

$$Y_0/G = \beta_0. \tag{4b}$$

The existence of undercoupling or overcoupling is determined experimentally by the presence of a minimum or a maximum, respectively, at the input window.

A calculating technique which makes somewhat better use of the data is as follows: The equation of the resonance curve is:

$$\left(\frac{2\Delta Q_0 G}{Y_0}\right)^2 \frac{Y_0}{G} + \frac{Y_0}{G} + \frac{G}{Y_0} = \frac{1+\beta^2}{\beta}, \tag{5}$$

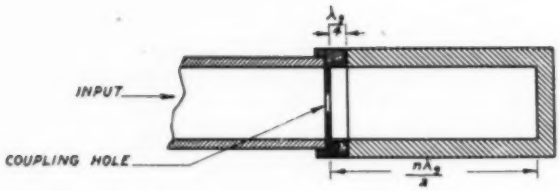


FIG. 3a. Two part resonator.

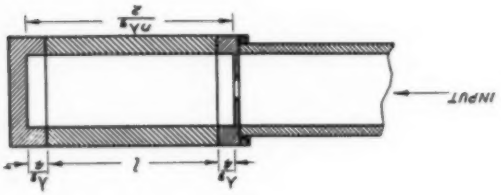


FIG. 3b. Three part resonator.

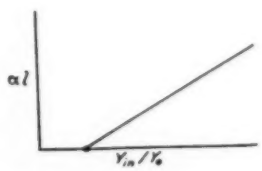


FIG. 3c. Calibration plot.

where $\Delta=(\omega-\omega_0/\omega_0)$, and $\beta=$voltage standing wave ratio at frequency ω. If one plots $(1+\beta^2/\beta)$ *vs.* Δ^2, a straight line is obtained. By fitting the best straight line to the experimental data, one may evaluate Q_0, Q_L, and Y_0/G. This procedure makes use of the entire resonance curve and not merely three points.

It is thus clear that one may determine Q_0 from bandwidth measurements and, having found it, calculate the corresponding effective conductivity of the wall material. In principle then, one may construct a series of cavities from the various materials to be tested and evaluate the effective conductivities by the methods described above.

Although this technique is straightforward, there were a number of practical reasons which militated against its use in this investigation. First of all, the active surfaces, being the interior walls of the cavity, would not always be readily accessible for polishing, plating, or other surface treatment. Secondly, the bandwidth measurements were tedious and difficult with the techniques then available, and since many measurements were to be made, they would have proved very time consuming. The bandwidth measurements require stable oscillators and means for measuring accurately small frequency intervals. Although such equipment is available today, it was not at the time this work was begun. Therefore, a modified technique was developed in which a minimum of bandwidth measurements was required.

Suppose we consider a cavity which is feeding power back to the line, instead of absorbing power from the line. This would be the case when the oscillator is shut off, and the field in the cavity starts to decay. Then,

$$Q_0=2\pi\frac{\text{Energy stored in the cavity}}{\text{Energy dissipated in cavity walls per cycle}}, \tag{6a}$$

$$Q_L=2\pi\frac{\text{Energy stored in the cavity}}{\text{Energy dissipated in cavity walls+energy radiated back to line per cycle}}, \tag{6b}$$

$$Q_W=2\pi\frac{\text{Energy stored in the cavity}}{\text{Energy radiated back to line per cycle}}. \tag{6c}$$

Q_W is the window Q, or external Q. It depends only on the shape of the cavity and the geometry of the coupling window and not upon the conductivity of the walls.

Evidently,

$$1/Q_L=(1/Q_0)+(1/Q_W). \tag{7}$$

Also since

$$Q_0/Q_L=1+(Y_0/G), \quad Y_0/G=Q_0/Q_W. \tag{8}$$

Equation (8) means that the standing-wave ratio at resonance, or its reciprocal, is proportional to the unloaded Q.

Thus in principle, if we have a series of cavities which are identical in all dimensions, including those of the coupling window, the input standing-wave ratio of each cavity will be proportional to its Q, and the constant of proportionality is the same for all cavities. A bandwidth measurement on just one of the cavities will suffice to determine the window Q, and thus the constant of proportionality. The fact is, however, that the window Q varies very rapidly with the coupling window dimensions, so that it is impracticable to maintain the necessary tolerances.

A more practical approach is to build a cavity in two or more demountable sections, such that the test piece may form one of these sections. Examples of this type of construction are indicated in Figs. 3–6.

The first one tried was the TE_{011} resonator of Fig. 5a. The removable sample is the end plate. The lines of current flow are coaxial circles and nowhere cross the contact surface, so that the losses in the oscillating mode are not affected by the contact resistance. There is considerable advantage in having the sample in the form of a flat plate, since this is the most convenient form for cleaning, polishing, plating, etc. This resonator had two serious disadvantages, however. First of all, the TE_{011} mode is intrinsically degenerate with the TM_{111} odd and even modes, so that extreme care must be taken to avoid exciting these modes (if it can be avoided at all).

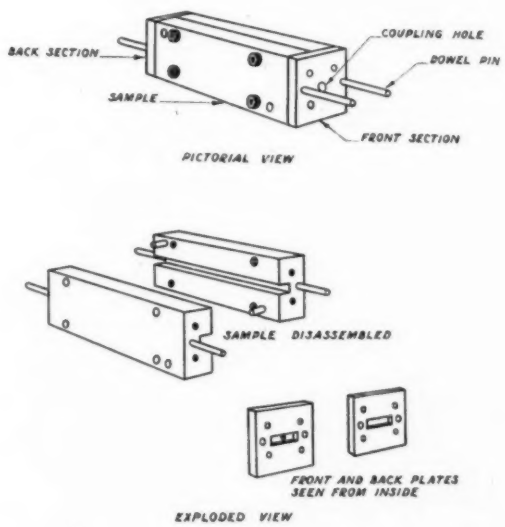


FIG. 4. Rectangular resonator details.

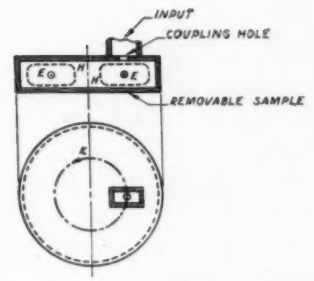


FIG. 5a. TE_{011} mode resonator.

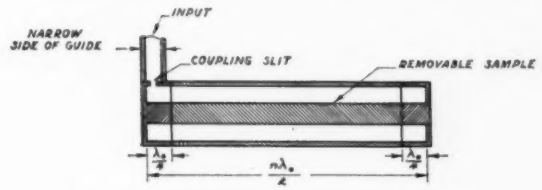


FIG. 5b. Coaxial resonator—TEM mode.

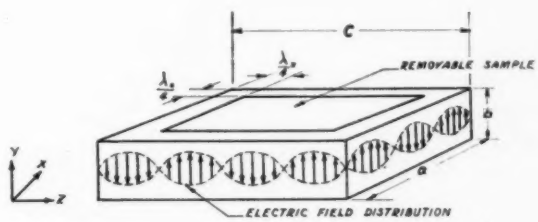


FIG. 6. Rectangular resonator.

Secondly, there are a number of non-resonant modes present in the cavity which influence the window Q in an undesirable fashion. This trouble seems to arise whenever we use a resonator whose dimensions are such that more than one mode can propagate in the cavity, even though the unwanted modes are non-resonant. For the cavity shown in Fig. 5a, the TE_{11}, TM_{01}, and TE_{21} modes, which are lower than the TE_{01}, will certainly propagate along the axis of the cylinder, and in addition many higher ones will also propagate because of the relatively large diameter. (A squat cavity was chosen in order to make the ratio of sample loss to total loss as large as possible.) These extraneous modes are excited because the presence of the coupling hole, or slit, imposes boundary conditions which cannot be satisfied by the resonant mode alone. The currents which flow, as a result of these modes, do cross the contact surfaces and are therefore controlled by the contact resistance.

The energy stored in these modes determines the equivalent susceptance of the coupling iris and hence the window Q. In other words, the window Q depends upon the contact resistance which is not a reproducible quantity. Therefore, the window Q varies from measurement to measurement and must be redetermined each time. Since it was desired to eliminate bandwidth measurements, this type of resonator was discarded.

There is, however, considerable merit in using flat plate samples. With the stabilized oscillators now available,[1] bandwidth measurements in the region of 24,000 Mc are no longer difficult so that the method of the preceding paragraph is entirely practicable today. The TE_{011} cavity is not the best choice, however, owing to the degeneracy with the TM_{111} modes. A suggested cavity, using a removable flat plate sample, is illustrated in Fig. 6. This would operate in the TE_{10n} mode. By choosing the dimensions such that $b<\lambda_0/2$, and if a and b are incommensurable, one may avoid accidental degeneracies. The resonant wave-length λ_0 is determined from the relation $(2/\lambda_0)^2=(l/a)^2+(n/c)^2$ in which l and n are equal to the number of half-period variations along the x and z axes, respectively. The contact surfaces are at current nodes. A cylindrical cavity in the TM_{0m0} mode could be used as well, in which case the sample plate would be circular.

Figure 5b illustrates another type of demountable cavity in which the sample is the removable center conductor. This resonator operates in the principal mode. The breaks in the inner and outer conductors occur at points of zero current. One version of this was tried in which the diameter was large enough so that the TE_{11} mode could propagate. Here again it was found that the window Q was not reproducible.

The cavity finally chosen for making most of the measurements was the demountable rectangular type illustrated in Figs. 3 and 4. This was taken as the best compromise between the conflicting requirements of convenient sample shape and simplicity of measuring technique. The cavity operates in the lowest mode, the TE_{10n}. The cross-sectional dimensions are standard waveguide dimensions for 1.25 cm, and therefore modes higher than the TE_{10} cannot propagate. The higher modes generated in the vicinity of the coupling window are damped out very rapidly. The breaks in the walls occur across lines of no current flow and are sufficiently removed from the neighborhood of the coupling window so that the higher mode amplitude is practically nil.

Both two- and three-section cavities were used, although the three-section type is preferred because it is easier to get at the active surfaces. An exploded view of the three-part resonator is shown in Fig. 4.

The normalized input conductance at resonance is:

$$G/Y_0=Q_w/Q_0. \tag{8a}$$

For the simple rectangular resonators used this may be written in the form

$$\frac{G}{Y_0}=\frac{n\pi B^2}{2Q_0}\left(\frac{\lambda_g}{\lambda_0}\right)^2=\left(\frac{B}{Y_0}\right)^2(\alpha l+R), \tag{9}$$

where n = the axial length in numbers of half wave-lengths, λ_g = the guide wave-length, λ_0 = the free space wave-length, B = the susceptance of the coupling window, α = the attenuation in the sample portion in nepers/meter, l = the length of the sample portion, and R = a resistance which represents the dissipation in the front and back quarter-wave sections.

In the above formula (αl) is the total attenuation in the sample. The relation between (αl) and G/Y_0 is linear. A convenient way to calibrate a pair of front and back quarter-wave sections is as follows. The attenuation in a long piece of waveguide is measured by the short-circuited guide technique described earlier in this paper. Samples with different values of (αl) are obtained by cutting this piece up into short sections, each one of which is a different multiple of a half wave-length. From the data taken with these samples a calibration plot of total attenuation in the sample *vs.* input conductance is obtained as in Fig. 3c.

This calibration plot is then used to determine the attenuation in unknown samples. The sample

[1] R. V. Pound, Radiation Laboratory Report 662; R. V. Pound, Radiation Laboratory Report 837; and R. V. Pound, "Electronic Frequency Stabilization of Microwave Oscillators," Rev. Sci. Inst. **17**, 490 (1946).

TABLE I. Conductor losses in standard K-band waveguide.[1]

Material	Measured atten. in db/meter	Calculated[2] atten. in db/meter	Meas. atten. / Calc. atten.	Eff. cond. at K-band in 10^7 mho/m	d.-c. Cond. in 10^7 mho/m
Aluminum[7]					
Pure, commercial (machined surface)	0.58	0.455	1.27	1.97	3.25 (measured)
17S Alloy[7] (machined surface)	0.75	0.586	1.28	1.19	1.95 (measured)
24S Alloy (machined surface)	0.66	0.635	1.04	1.54	1.66 (measured)
Brass					
Yellow (80–20) drawn wave guide	0.68	0.653	1.04	1.45	1.57 (measured)
Red (85–15) drawn wave guide	0.55			2.22	
Yellow round drawn tubing[3]	0.90	0.844	1.07	1.36	1.56 (measured)
Yellow (80–20) (machined surface)	0.76	0.653	1.15	1.17	1.57 (measured)
Free machining brass (mach. surface)	0.75	0.673	1.12	1.11	1.48 (measured)
Cadmium plate	0.79–0.87	0.711	1.13–1.22	1.04–0.89	1.33 Hdbk. of Phys. and Chem.
Chromium plate, dull[4]	0.67–0.82	0.418	1.60–1.96	1.49–0.99	3.84 Hdbk. of Phys. and Chem.
Copper					
Drawn O.F.C. wave guide	0.41	0.350	1.17	4.00	5.48 (measured)
Drawn round tubing	0.52	0.496	1.05	4.10	4.50 (measured)
Machined surface[7]	0.38	0.349	1.09	4.65	5.50 (measured)
Copper plate	0.54–0.61	0.337	1.60–1.81	2.28–1.81	5.92 } Hdbk. of Phys. and Chem.
Electroformed wave guide[7]	0.46	0.337	1.37	3.15	5.92 }
Gold plate	0.60	0.404	1.48	1.87	4.10 Hdbk. of Phys. and Chem.
Iron, electroformed[7]	1.01				
Mercury[6]	2.50	2.54	0.98	0.104	0.104 Hdbk. of Phys. and Chem.
Monel (machined surface)[5,7]	2.08	2.07	1.01	0.155	0.156 (measured)
Nickel[7]					
Electroformed wave guide	0.82				
Nickel plate	1.11				
Silver					
Coin silver drawn wave guide	0.45	0.375	1.20	3.33	4.79 (measured)
Coin silver lined wave guide	0.60	0.375	1.60	1.87 }	4.79 (assumed)
Coin silver (machined surface)[7]	0.51	0.375	1.34	2.66 }	
Fine silver (machined surface)[7]	0.48	0.330	1.45	2.92 }	6.14 Hdbk. of Phys. and Chem.
Silver plate	0.41–0.57	0.330	1.24–1.73	3.98–2.05 }	
Solder, soft[7]	1.05	0.978	1.08	.600	0.70 (measured)
Steel, cold rolled (mach. surface)	2.85				

[1] Unless otherwise noted, figures are for TE_{10} mode, $\lambda = 1.25$ cm in rectangular guide of dimensions of 0.170″ ×0.420″.
[2] Theoretical attenuation for TE_{10} mode in 0.170″ ×0.420″ rectangular waveguide at 1.25 cm is $2590\sigma^{-\frac{1}{2}}$ db/m where σ is given in mho/meters. For the TE_{11} mode in round pipe at 1.25 cm, the attenuation is $3330\sigma^{-\frac{1}{2}}$ db/m.
[3] 0.345″ I.D., TE_{11} mode.
[4] No nickel undercoat.
[5] This surface was somewhat rougher than most machined surfaces.
[6] For method of measurement refer to the Appendix. Figures are expressed for a guide of 0.170″ ×0.420″ cross section.
[7] Only one sample of these was tested.

may be any multiple of a half wave-length. However, for any given window susceptance there is a range of sample attenuations which produce conveniently measurable standing-wave ratios. It is a good idea to plate the front and back sections with some corrosion-resistant material to insure permanence of calibration.

3. EXPERIMENTAL RESULTS

The more important available data on conductor losses at 1.25 cm are summarized in Table I. These are expressed in terms of attenuation of the TE_{10} mode in standard 0.170″ ×0.420″ waveguide at 1.25 cm, or in some cases for the TE_{11} mode in round pipe of 0.345″ diameter. The first column records the value of the experimentally measured attenuation, the second that of the theoretically computed attenuation, and the third column gives the ratio of the two; the fourth and fifth columns list the effective conductivity at 1.25 cm and the d.c. conductivity, respectively.

The theoretical attenuations are computed by means of the formulas given in the footnotes of Table I. The conductivities used in these formulas are the d.c. conductivities listed in the fifth column. The d.c. conductivity of the samples was measured wherever possible but in some cases it was not convenient or possible to do so. In these other cases, the sources from which the d.c. conductivity data were taken, are indicated. In comparing theoretical and experimental performance it is important to have accurate data on the d.c. conductivities if the comparison is to have much significance. In the case of fairly pure materials, such as mercury or electrodeposited metals, the conductivity figures given in handbooks for pure metals are probably good enough, but for commercial metals and alloys there may be discrepancies between tabulated and actual values.

Although the listing is far from complete, many of the common conductor materials are present. There is usually some variation in con-

ductivity among different specimens of the same material, particularly in machined and electroplated surfaces. Where the variation among specimens was more than a few percent, the outside limits of attenuation are stated in the table. Some of the figures are based on a single specimen and are so indicated. In the other cases anywhere from two to about six samples were tested. The probable error to be associated with the attenuation figures is estimated as less than two percent. In the case of the soft solder the error may be of the order of 10 percent since the measurement quoted here is of an earlier vintage than the others and the technique employed was less accurate.

The surfaces examined fall into three categories: drawn, machined, and electrodeposited. The drawn surfaces are as a rule quite good and although their losses are greater than theory predicts, the discrepancy is not large. In the case of brass it may amount to only 5 percent. Machined surfaces are frequently poorer, while plated surfaces seem to vary a good deal.

It seems reasonable to assume that the increase in r-f conductivity over the d.c. value is due primarily to surface roughness. It is difficult, however, to establish such a correlation on the basis of the observed data since no reliable index of surface roughness is available for the samples tested.

The measurement of the r-f conductivity of mercury, which was briefly mentioned earlier, is of interest in this connection. In this experiment the attenuation was measured in a waveguide made by immersing a smooth polystyrene rod in mercury. Because of the surface tension of the mercury, the resulting surface was considered to be free from the usual multitude of scratches and crevices present on solid metallic surfaces. Furthermore, since the skin depth is about eight times as big in mercury as in copper or silver, such irregularities as remain are relatively less important. The measured attenuation agreed with the calculated value within 2 percent which was about as accurately as the standing-wave ratio could be measured. Actually, however, the experimental figure may be in doubt by 4 or 5 percent as a result of a possible uncertainty in the assumed value of the conductivity of coin silver which enters into the experiment. The fact that agreement was attained here between theory and experiment again implies that the lack of agreement observed in other cases is due to surface irregularity. This matter is discussed in the Appendix where the experiment is described in detail.

TABLE II. Conductor losses in K-band waveguide with protective coatings.

Material	Measured attenuation in db/meter
Palladium flash (10^{-5} in.) on coin silver	0.6
Rhodium flash (10^{-5} in.) on coin silver	1.0
Sperry f710 lacquer on copper plate	1.2
Same surface without lacquer	0.6
Zinc chromate olive drab primer on 17S aluminum alloy	1.0

It should be noted that for the monel sample also, there is substantial agreement between observed and calculated attenuations. The skin depth in monel is approximately the same as in mercury. Thus one may infer that because of the greater skin depth the scratches are less important.

Many attempts were made to improve the conductivity of samples by polishing the active surfaces by both mechanical and electrolytic methods, but none of these were successful. Probably something of the order of a metallographic or optical polish is necessary, but it was not possible to approach this degree of excellency on account of the awkward shape of the surfaces to be polished. In this connection, flat samples would be decidedly advantageous.

The reverse process was possible. The conductivity was definitely lowered in many cases by abrading the surface. For example, when some yellow brass waveguide tubing was broached out, the attenuation was increased from 0.68 db/m to 0.74 db/m.

Thin oxide films are not harmful as long as the oxide resistivity is high. It may be shown that the effect of oxide films is small as long as the skin depth in the oxide material is large compared to the oxide layer thickness.

The effect of a few protective coatings in increasing attenuation is shown in Table II.

Of all the solid metals listed mild steel has the greatest loss. This behavior has been explained

by Kittel on the basis of the magnetic properties.[2]

ACKNOWLEDGMENT

I am indebted to Dr. E. M. Purcell, formerly of the Radiation Laboratory, for many helpful suggestions.

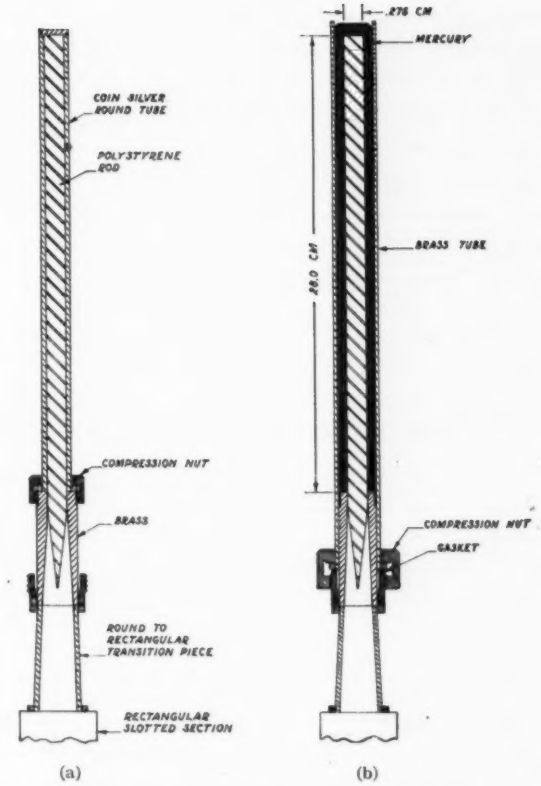
APPENDIX

The Mercury Waveguide Experiment

The experimental setup is indicated in Figs. 7a and 7b. The wave guide consisted of a round polystyrene rod immersed in mercury. This was tapered into a round air-filled guide as shown in Fig. 7b. A round-to-rectangular transition coupled the round guide to a standard rectangular slotted section. The rod was then treated as a shorted round waveguide and the attenuation measured by the techniques described under method I for round guide.

The dielectric losses were separated from the conductor losses in the following manner. Before setting up the mercury guide a tightly fitting drawn coin silver tube was slipped over the polystyrene rod, as in Fig. 7a. The input standing-wave ratio was measured, and from this the total attenuation in the polystyrene-silver guide was calculated. The silver tube was then removed and the polystyrene rod immersed in mercury as in Fig. 7b. The standing-wave ratio was again measured and the attenuation in the polystyrene system calculated. The difference between these two figures is the conductor attenuation in the mercury minus the conductor attenuation in the coin silver. The attenuation in the coin silver was calculated by using the effective conductivity established from previous measurements on drawn rectangular coin silver wave guide, and the attenuation due to the mercury was determined. The attenuation thus found agreed with theory within 2 percent.

An estimate of the limit of error was made on the following basis. The attenuation in the polystyrene-silver system was 1.46 db; in the polystyrene-mercury system it was 3.29 db, mercury attenuation less silver attenuation



FIGS. 7a and 7b. Mercury wave guide.

$=1.83$ db. The effective conductivity of coin silver wave guide is 3.3×10^7 mhos per meter. The conductor attenuation for the TE_{11} mode in the silver tube was 0.40 db for a 28 cm length (see Fig. 7b). The net loss due to the mercury alone was therefore 2.23 db or 8.0 db per meter. The value calculated by using the tabulated figure of 0.1044×10^7 mhos per meter for mercury at 20°C, was 8.10 db per meter. The error depends largely on the accuracy of the assumed attenuation for the coin silver. By assuming a 20 percent error in estimating the attenuation in silver, which should certainly be an outside figure, the uncertainty in the mercury attenuation would be less than 4 percent.

It should also be pointed out that residual mismatch in the taper or transition causes very little error because of the fact that the mismatch appears in both measurements and almost completely cancels out in the final result. As a preliminary part of the experiment, taper and transition mismatch was investigated by slipping

[2] C. Kittel, "Theory of the dispersion of magnetic permeability in ferromagnetic materials at microwave frequencies," Phys. Rev. **70**, 281 (1946).

some close fitting poly-iron sleeves over the polystyrene rod, in place of the silver tube, so as to obtain a matched absorbing load. The voltage standing-wave ratio looking in was 1.08. It may be verified by calculation that in the worst case this would make an error of less than ½ percent.

Formula Used in Computing Attenuation:[3]

$$\alpha=\frac{1}{a}\left(\frac{\pi\epsilon\nu}{\sigma}\right)^{\frac{1}{2}}\left[1-\left(\frac{\nu_{11}'}{\nu}\right)^2\right]^{-\frac{1}{2}}\times\left[\left(\frac{\nu_{11}'}{\nu}\right)^2+\frac{1/u_{11}'^2}{1-1/u_{11}'^2}\right].$$

a = radius of polystyrene rod, $\epsilon=2.52\epsilon_0$ for polystyrene, $u_{11}'=1.84$, $\nu_{11}'=[u_{11}'/2\pi a(\epsilon\mu)^{\frac{1}{2}}]$, and ν = frequency. M.K.S. units used.

[3] See J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., 1941), p. 544.

Note on Circular Loop Antennas with Non-Uniform Current Distribution

G. Glinski
Northern Electric Company, Montreal, Quebec, Canada
(Received February 6, 1947)

In this paper are presented the approximate formulas for the components of radiation vectors of a short-circuited circular loop with non-uniform current distribution. The formulas are valid for the ratio of loop perimeter to wavelength of the order of 0.5 or less, and assume the current distribution of the hyperbolic cosine form. These formulas lead to the radiation intensity formula from which the expressions for the horizontal and vertical field patterns are derived. The latter formulas are further simplified assuming that the attenuation constant is much smaller than the phase constant. From the expression for horizontal field pattern, it follows that the pattern is symmetrical about the loop axis of symmetry. Moreover, this horizontal pattern exhibits a directional effect with maximum field in the direction of the feeder end of the loop. This directional effect is a function of loop dimensions. The theoretical horizontal pattern agrees very closely with the experimental one. From the expression for vertical field pattern, it follows that the non-uniform current distribution produces a pattern intermediate between that for horizontal dipole and horizontal small loop with uniform current distribution. The field intensity in the zenithal direction is again a function of loop dimensions. Using the expression for radiation intensity, formulas for a radius of equivalent circular horizontal field pattern, power gain, average power gain, and radiation resistance are derived. It is shown that the average power gain is essentially a function of loop radius and decreases with the increase of the latter. Finally, the approximate expression for the attenuation constant of the transmission line equivalent to the loop is derived.

THE purpose of this note is to present formulas for radiation pattern and gain of circular loop antennas with non-uniform current distribution. Although the problem is of considerable practical importance for the designer of some f-m broadcast antennas[1–4] whose elements are often circular loops or may be considered as such to the first approximation, there appears to be very little in the literature on the subject.[5,6]

The general problem of current distribution on a circular radiating loop has been treated previously.[7,8]

In the following, the actual current distribution will be approximated by that on the equivalent uniform transmission line.[9] The effect of radiation will be included in the attenuation of this line.[10]

GENERAL FORMULA FOR RADIATION VECTOR OF CIRCULAR LOOP

Consider the circular loop in $(\theta=\pi/2)$-plane of the spherical coordinate system r, φ, θ of

[1] N. W. Scheldorf, Gen. Elec. Rev. **46**, 163 (1943).
[2] C. R. Jones, Communications **26**, 36 (1946).
[3] Bell Lab. Record **24**, 163 (1946).
[4] R. F. Holz, F.M. and Television **6**, 45 (1946).
[5] E. M. Williams, Proc. I.R.E. **28**, 480 (1940).
[6] J. B. Sherman, Proc. I.R.E. **32**, 534 (1944).
[7] E. Hallen, Nova Acta Regia Soc. Sci. Upsaliensis, Series IV, **11**, 33 (1938).
[8] F. B. Pidduck, *Currents in Aerials and High-Frequency Networks* (Clarenden Press, Oxford, 1946), p. 62.
[9] L. L. Libby, Proc. I.R.E. **34**, 641 (1946).
[10] S. A. Schelkunoff, *Electromagnetic Waves* (D. Van Nostrand Company, Inc., New York, 1943), p. 200.

Fig. 1. The center of the loop is at the origin of the coordinate system.

Consider two diametrically opposite loop elements A and B with the coordinates $\bar{r}$, $\bar{\varphi}$, $\pi/2$, and $\bar{r}$, $\bar{\varphi}+\pi$, $\pi/2$, and currents of complex amplitudes I_A and I_B, respectively.

Following Schelkunoff[11] the components of elementary radiation vectors are

$$\begin{aligned} dN_{A,x} &= -I_A\bar{r}\sin\bar{\varphi}d\bar{\varphi}e^{jt}, \\ dN_{A,y} &= I_A\bar{r}\cos\bar{\varphi}d\bar{\varphi}e^{jt}, \\ dN_{B,x} &= I_B\bar{r}\sin\bar{\varphi}d\bar{\varphi}e^{-jt}, \\ dN_{B,y} &= -I_B\bar{r}\cos\bar{\varphi}d\bar{\varphi}e^{-jt}, \end{aligned} \tag{1}$$

where $t=\beta\bar{r}\sin\theta\cos(\varphi-\bar{\varphi})$.

The components of equivalent elementary radiation vector are therefore

$$\begin{aligned} dN_x &= -I_1\bar{r}\sin\bar{\varphi}d\bar{\varphi}, \\ dN_y &= I_1\bar{r}\cos\bar{\varphi}d\bar{\varphi}, \end{aligned} \tag{2}$$

where $I_1=I_Ae^{jt}-I_Be^{-jt}$.

It is convenient to introduce the symmetrical and anti-symmetrical components of currents defined by

$$\begin{aligned} I_A &= (I_s+I_a)/2, \\ I_B &= (I_s-I_a)/2. \end{aligned} \tag{3}$$

In terms of I_s and I_a, formulas (2) can be written as:

$$\begin{aligned} dN_x &= -I_1\bar{r}\sin\bar{\varphi}d\bar{\varphi}, \\ dN_y &= I_1\bar{r}\cos\bar{\varphi}d\bar{\varphi}, \end{aligned} \tag{4}$$

where $I_1=jI_s\sin t+I_a\cos t$.

In the case of f-m broadcast antennas, $\beta\bar{r}$ is of the order of 0.5. It is quite accurate therefore to replace the trigonometric functions in formula (4) by first terms of their Bessel expansion[12] (Appendix 1). To the first approximation therefore

$$\begin{aligned} dN_x &\cong -I_2\bar{r}\sin\bar{\varphi}d\bar{\varphi}, \\ dN_y &\cong I_2\bar{r}\cos\bar{\varphi}d\bar{\varphi}, \end{aligned} \tag{5}$$

where $I_2=J_0(\beta\bar{r}/2)[j4I_sJ_1(\beta\bar{r}/2)\sin\theta\cos(\varphi-\bar{\varphi})+I_aJ_0(\beta\bar{r}/2)]$.

The components of total radiation vector are found by summing up the contributions of all elementary components of radiation vectors:

$$N_x=\int_0^{\pi}dN_x, \quad N_y=\int_0^{\pi}dN_y. \tag{6}$$

[11] S. A. Schelkunoff, reference 10, p. 336.
[12] S. A. Schelkunoff, reference 10, p. 55.

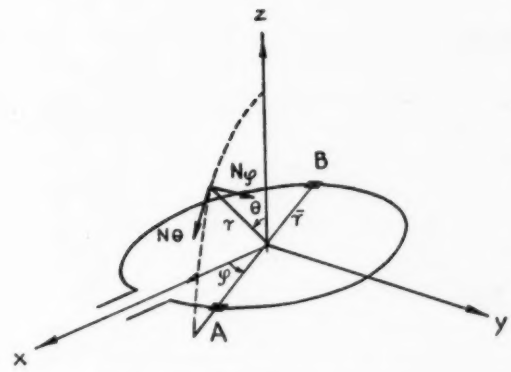


FIG. 1. Coordinate system used.

RADIATION VECTOR OF SHORT-CIRCUITED CIRCULAR LOOP

Consider the important practical case of the loop fed at $\bar{\varphi}=0$, and short-circuited at $\bar{\varphi}=\pi$. The current distribution on the equivalent transmission line is

$$I_l=I\cosh\gamma l; \tag{7}$$

where I = current at the short-circuit; γ = propagation constant $=\alpha+j\beta$; l = distance from the short-circuit. I_A and I_B of formulas (3) are:

$$\begin{aligned} I_A &= I\cosh\gamma\bar{r}(\pi-\bar{\varphi}), \\ I_B &= I\cosh\gamma\bar{r}\bar{\varphi}. \end{aligned} \tag{8}$$

Therefore,

$$\begin{aligned} I_s &= I[(\cosh\gamma\bar{r}\pi+1)\cosh\gamma\bar{r}\bar{\varphi} \\ &\qquad -\sinh\gamma\bar{r}\pi\sinh\gamma\bar{r}\bar{\varphi}], \\ I_a &= I[(\cosh\gamma\bar{r}\pi-1)\cosh\gamma\bar{r}\bar{\varphi} \\ &\qquad -\sinh\gamma\bar{r}\pi\sinh\gamma\bar{r}\bar{\varphi}]. \end{aligned} \tag{9}$$

The elementary integration and reduction (Appendix 2) of formulas (6), using the values of I_s and I_a from formula (9), gives

$$N_x=I\bar{r}J_0\left[-jJ_1\sin\theta\frac{16\sin j\gamma\bar{r}\pi}{j\gamma\bar{r}(4+\gamma^2\bar{r}^2)}\sin\varphi\right], \tag{10}$$

$$N_y=I\bar{r}J_0\left[jJ_1\sin\theta\frac{8(2+\gamma^2\bar{r}^2)\sin j\gamma\bar{r}\pi}{j\gamma\bar{r}(4+\gamma^2\bar{r}^2)}\cos\varphi - J_0\frac{2j\gamma\bar{r}\sin j\gamma\bar{r}\pi}{1+\gamma^2\bar{r}^2}\right], \tag{11}$$

where $J_0\equiv J_0(\beta\bar{r}/2)$, $J_1\equiv J_1(\beta\bar{r}/2)$.

The spherical components of total radiation vector are, therefore,[13]

[13] S. A. Schelkunoff, reference 10, p. 234.

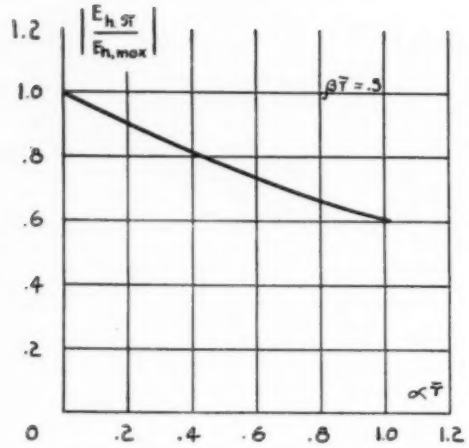


FIG. 2. Ratio of electrical field intensities in ($\varphi=\pi$)- and ($\varphi=0$)-directions in the plane of the loop *versus* attenuation constant $\alpha\bar{r}$ for $\beta\bar{r}=0.5$.

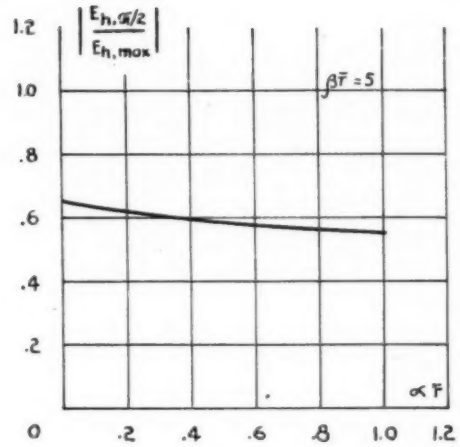


FIG. 3. Ratio of electrical field intensities in ($\varphi=\pi/2$)- (or $3\pi/2$) and ($\varphi=0$)-directions in the plane of the loop *versus* attenuation constant $\alpha\bar{r}$ for $\beta\bar{r}=0.5$.

$$N_\theta=(N_x\cos\varphi+N_y\sin\varphi)\cos\theta$$

$$=A\left[-j\frac{J_1}{J_0}\frac{\sin2\theta\sin2\varphi}{4+\gamma^2\bar{r}^2}-\frac{\cos\theta\sin\varphi}{1+\gamma^2\bar{r}^2}\right]; \quad (12)$$

where $A=2I\bar{r}J_0^2j\gamma\bar{r}\sin j\gamma\bar{r}\pi$,

$$N_\varphi=-N_x\sin\varphi+N_y\cos\varphi$$

$$=A\left[-j\frac{J_1}{J_0}\frac{4(2+\gamma^2\bar{r}^2\cos^2\varphi)\sin\theta}{\gamma^2\bar{r}^2(4+\gamma^2\bar{r}^2)}-\frac{\cos\varphi}{1+\gamma^2\bar{r}^2}\right]. \quad (13)$$

Using the approximations of Appendix 3, the formulas (12) and (13) reduce to:

$$N_\theta=A[(-f(J_1/J_0)\sin2\theta\sin2\varphi-a\cos\theta\sin\varphi)+j(-e(J_1/J_0)\sin2\theta\sin2\varphi+b\cos\theta\sin\varphi)], \quad (14)$$

$$N_\varphi=A[(-8d(J_1/J_0)\sin\theta-4f(J_1/J_0)\sin\theta\cos^2\varphi-a\cos\varphi)+j(-8c(J_1/J_0)\sin\theta-4e(J_1/J_0)\sin\theta\cos^2\varphi+b\cos\varphi)]. \quad (15)$$

FIELD PATTERN OF SHORT-CIRCUITED CIRCULAR LOOP

The radiation intensity[14] is

$$\Phi=\frac{15\pi}{\lambda^2}(N_\theta N_\theta^*+N_\varphi N_\varphi^*), \quad (16)$$

where asterisk denotes conjugate quantity. Therefore, using formulas (14) and (15)

$$\Phi=\frac{15\pi|A|^2}{\lambda^2}\left[\left(f\frac{J_1}{J_0}\sin2\theta\sin2\varphi+a\cos\theta\sin\varphi\right)^2+\left(e\frac{J_1}{J_0}\sin2\theta\sin2\varphi+b\cos\theta\sin\varphi\right)^2+\left(8d\frac{J_1}{J_0}\sin\theta+4f\frac{J_1}{J_0}\sin\theta\cos^2\varphi+a\cos\varphi\right)^2+\left(8c\frac{J_1}{J_0}\sin\theta+4e\frac{J_1}{J_0}\sin\theta\cos^2\varphi-b\cos\varphi\right)^2\right]. \quad (17)$$

In practical antennas $d\gg f$ and $c\gg e$; therefore, neglecting the terms of the order of f and e in the last two round brackets:

$$\Phi=\frac{15\pi|A|^2}{\lambda^2}\left[\left(f\frac{J_1}{J_0}\sin2\theta\sin2\varphi+a\cos\theta\sin\varphi\right)^2+\left(e\frac{J_1}{J_0}\sin2\theta\sin2\varphi+b\cos\theta\sin\varphi\right)^2+\left(8d\frac{J_1}{J_0}\sin\theta+a\cos\varphi\right)^2+\left(8c\frac{J_1}{J_0}\sin\theta-b\cos\varphi\right)^2\right]. \quad (18)$$

The field intensity[15] is $|E|=(240\pi\Phi)^{\frac{1}{2}}/r$. Using

[14] S. A. Schelkunoff, reference 10, p. 235.

[15] S. A. Schelkunoff, reference 10, p. 333.

formula (18), the horizontal field pattern is:

$$|E_h| = B\left[\left(8d\frac{J_1}{J_0}+a\cos\varphi\right)^2+\left(8c\frac{J_1}{J_0}-b\cos\varphi\right)^2\right]^{\frac{1}{2}}, \tag{19}$$

where $B=60\pi|A|/r\lambda$.

From the definition of the coefficients a, b, c, and d, it follows that: $a>0$, $b>0$, $c<0$, $d>0$. Therefore, the maximum field intensity in the horizontal plane is for $\varphi=0$

$$|E_{h,\max}| = B\left[\left(8d\frac{J_1}{J_0}+a\right)^2+\left(8c\frac{J_1}{J_0}-b\right)^2\right]^{\frac{1}{2}}. \tag{20}$$

The field intensity in the horizontal plane for $\varphi=\pi$ is

$$|E_{h,\pi}| = B[(8d(J_1/J_0)-a)^2 + (8c(J_1/J_0)+b)^2]^{\frac{1}{2}}. \tag{21}$$

The ratio of field intensities in $\varphi=\pi$ and $\varphi=0$ directions is, therefore,

$$|E_{h,\pi}|/|E_{h,\max}| = \left[\frac{(8d(J_1/J_0)-a)^2+(8c(J_1/J_0)+b)^2}{(8d(J_1/J_0)+a)^2+(8c(J_1/J_0)-b)^2}\right]^{\frac{1}{2}}. \tag{22}$$

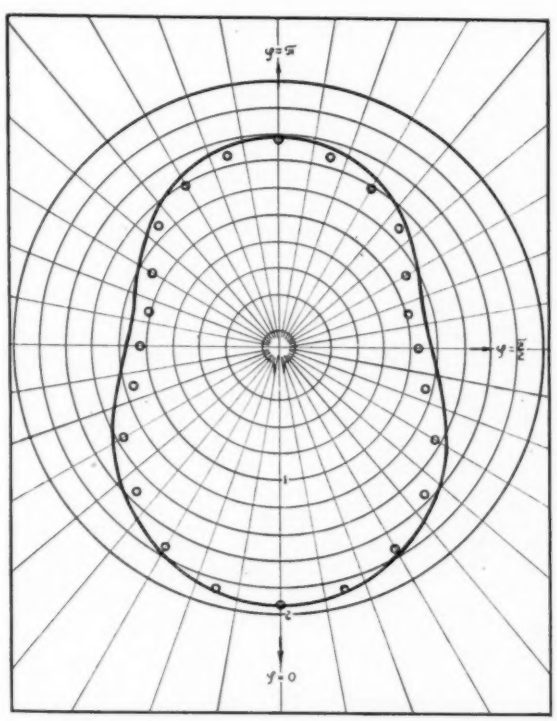


FIG. 4. Horizontal field pattern for $\alpha\bar{r}=0.04$ and $\beta\bar{r}=0.5$. Encircled points were obtained experimentally.

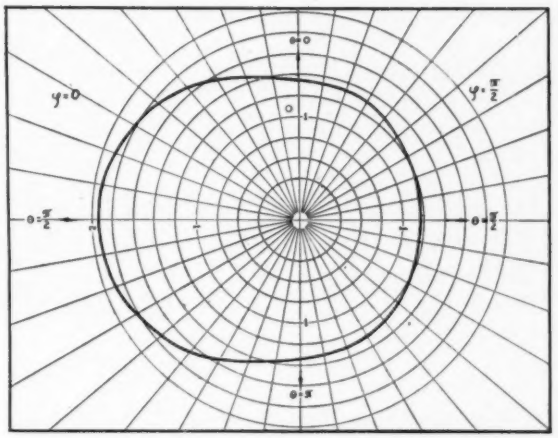


FIG. 5. Vertical field pattern in ($\varphi=0$)-(left) and ($\varphi=\pi/2$)-(right) planes for $\alpha\bar{r}=0.04$ and $\beta\bar{r}=0.5$.

From the last formula and the definition of coefficients a, b, c, d, it follows that the ratio of field intensities in $\varphi=\pi$ and $\varphi=0$ directions is a function of $\alpha\bar{r}$ and $\beta\bar{r}$. Figure 2 shows this ratio *versus* $\alpha\bar{r}$ for the important practical case of $\beta\bar{r}=0.5$.

The field intensity in the horizontal plane for $\varphi=\pi/2$ (or $3\pi/2$) is

$$|E_{h,\pi/2}| = B\left[\left(8d\frac{J_1}{J_0}\right)^2+\left(8c\frac{J_1}{J_0}\right)^2\right]^{\frac{1}{2}}. \tag{23}$$

The ratio of field intensities in $\varphi=\pi/2$ and $\varphi=0$ directions is

$$|E_{h,\pi/2}|/|E_{h,\max}| = \left[\frac{(8d(J_1/J_0))^2+(8c(J_1/J_0))^2}{(8d(J_1/J_0)+a)^2+(8c(J_1/J_0)-b)^2}\right]^{\frac{1}{2}}. \tag{24}$$

This ratio is again a function of $\alpha\bar{r}$ and $\beta\bar{r}$. Figure 3 shows this ratio *versus* $\alpha\bar{r}$ for $\beta\bar{r}=0.5$. Figure 4 is the plot of horizontal field pattern according to formula (18) for $\alpha\bar{r}=0.04$, and $\beta\bar{r}=0.5$. The encircled points represent the experimental pattern plotted to the same scale, and having the same value of $|E_{h,\pi}|/|E_{h,\max}|$ as the theoretical one.

It can be concluded that the horizontal field pattern of the circular short-circuited loop has the following characteristics:

1. The pattern is symmetrical about the $(0, \pi)$-direction.

2. There is the maximum field intensity in ($\varphi=0$)-direction (feeder end).

The vertical field pattern is

$$|E_v| = B[(f(J_1/J_0)\sin 2\theta \sin 2\varphi + a\cos\theta\sin\varphi)^2 + (e(J_1/J_0)\sin 2\theta \sin 2\varphi + b\cos\theta\sin\varphi)^2 + (8d(J_1/J_0)\sin\theta + a\cos\varphi)^2 + (8c(J_1/J_0)\sin\theta - b\cos\varphi)^2]^{\frac{1}{2}}. \quad (25)$$

In any φ-direction, the field intensity has a maximum value for $\theta = \pi/2$, that is in the horizontal plane.

The zenithal value of the field intensity ($\theta = 0$) is

$$|E_{v,z}| = B(a^2+b^2)^{\frac{1}{2}}, \quad (26)$$

and is independent of φ.

Figure 5 is the plot of vertical field pattern according to formula (24) for $\alpha\bar{r} = 0.04$ and $\beta\bar{r} = 0.5$, for $\varphi = 0$ and $\varphi = \pi/2$.

RADIUS OF EQUIVALENT CIRCULAR HORIZONTAL FIELD PATTERN

The field intensity of the equivalent circular horizontal field pattern can be found from the formula

$$\pi |E_{h,hv}|^2 = \tfrac{1}{2}\int_0^{2\pi} |E_h|^2 d\varphi. \quad (27)$$

Using formula (19) and after the elementary integration (Appendix 4)

$$|E_{h,hv}| = B[64(J_1/J_0)^2(d^2+c^2) + 0.5(a^2+b^2)]^{\frac{1}{2}}. \quad (28)$$

POWER GAIN OF SHORT-CIRCUITED CIRCULAR LOOP

The radiation intensity from formula (18), when reduced to the unity maximum, is

$$\Phi^1 = \frac{\Phi}{\Phi_{\theta=\pi/2,\,\varphi=0}} = \frac{1}{64(J_1/J_0)^2(d^2+c^2) + 16(J_1/J_0)(ad-bc) + (a^2+b^2)}[(J_1/J_0)^2(e^2+f^2)\sin^2 2\theta \sin^2 2\varphi + 2(J_1/J_0)(af+be)\sin 2\theta\cos\theta\sin 2\varphi\sin\varphi + (a^2+b^2)\cos^2\theta\sin^2\varphi + 64(J_1/J_0)^2(d^2+c^2)\sin^2\theta + 16(J_1/J_0)(ad-bc)\cos\varphi\sin\theta + (a^2+b^2)\cos^2\varphi]. \quad (29)$$

The radiated power[16] is

$$P = \int_{\theta=0}^{\theta=\pi}\int_{\varphi=0}^{\varphi=2\pi} \Phi^1 \sin\theta d\theta d\varphi. \quad (30)$$

Therefore (Appendix 5):

$$P = \frac{(8\pi/3)\{(J_1/J_0)^2[0.4(e^2+f^2) + 64(d^2+c^2)] + (a^2+b^2)\}}{64(J_1/J_0)^2(d^2+c^2) + 16(J_1/J_0)(ad-bc) + (a^2+b^2)}. \quad (31)$$

If short dipole is used as a reference,[17] the power gain is

$$(PG) = \frac{(8\pi/3)}{P} = \frac{64(J_1/J_0)^2(d^2+c^2) + 16(J_1/J_0)(ad-bc) + (a^2+b^2)}{(J_1/J_0)^2[0.4(e^2+f^2) + 64(d^2+c^2)] + (a^2+b^2)}. \quad (32)$$

AVERAGE POWER GAIN

In f-m broadcasting, it is customary to base the gain on the equivalent circular horizontal field pattern. Since the average radiation intensity in the horizontal plane is

$$\Phi_{hv} = r^2|E_{h,hv}|^2/240\pi = (r^2B^2/240\pi)[64(J_1/J_0)^2(d^2+c^2) + 0.5(a^2+b^2)], \quad (33)$$

and the maximum radiation intensity in the horizontal plane is

$$\Phi_{\max} = \Phi_{\theta=\pi/2,\varphi=0} = \frac{r^2B^2}{240\pi}\left[\left(8d\frac{J_1}{J_0}+a\right)^2 + \left(8c\frac{J_1}{J_0}-b\right)^2\right]; \quad (34)$$

therefore, the average power gain is

$$(PG)_{hv} = \frac{\Phi_{hv}}{\Phi_{\max}}(PG) = \frac{64(J_1/J_0)^2(d^2+c^2) + 0.5(a^2+b^2)}{(J_1/J_0)^2[0.4(e^2+f^2) + 64(d^2+c^2)] + (a^2+b^2)}. \quad (35)$$

[16] S. A. Schelkunoff, reference 10, p. 333.
[17] S. A. Schelkunoff, reference 10, p. 337.

The graph of $(PG)_{av}$ *versus* $\beta\bar{r}$ for $\alpha\bar{r}=0.04$ is shown in Fig. 6. (Since $c\gg d$ and $a\gg b$ in the practical range of $\alpha\bar{r}$, the gain computed for $\alpha\bar{r}=0$ may be applied also to the general case when $\alpha\bar{r}\neq 0$.)

TRANSMISSION LINE EQUIVALENT TO THE CIRCULAR LOOP WITH NON-UNIFORM CURRENT DISTRIBUTION

The radiated power P is given by formula (31). If the maximum current (at $\varphi=\pi$) is I, the radiation resistance R_r is

$$R_r=\frac{2P}{I^2}=\frac{16\pi}{3I^2}\cdot\frac{(J_1/J_0)^2[0.4(e^2+f^2)+64(d^2+c^2)]+(a^2+b^2)}{64(J_1/J_0)^2(d^2+c^2)+16(J_1/J_0)(ad-bc)+(a^2+b^2)}. \tag{36}$$

Since P in formula (31) was calculated for maximum radiation intensity equal to unity, therefore, from formula (18):

$$I^2=1\Big/\frac{15\pi|A|^2}{\lambda^2}\left[\left(8d\frac{J_1}{J_0}+a\right)^2+\left(8c\frac{J_1}{J_0}-b\right)^2\right]. \tag{37}$$

$$R_r=\frac{80\pi^2|A|^2}{\lambda^2}\left\{\left(\frac{J_1}{J_0}\right)^2[0.4(e^2+f^2)+64(d^2+c^2)]+(a^2+b^2)\right\}. \tag{38}$$

To calculate the attenuation constant α of the equivalent transmission line, let its resistance per unit length be ρ, and the current

$$I_{\bar{\varphi}}=I\cos\beta\bar{r}(\pi-\bar{\varphi}).$$

The power dissipated in the entire line is

$$P_d=\int^{\pi}\tfrac{1}{2}I_{\bar{\varphi}}^2\rho\bar{r}\,d\bar{\varphi}$$

$$=\tfrac{1}{4}I^2\rho\bar{r}\pi[1+(\sin 2\beta\bar{r}\pi/2\beta\bar{r}\pi)]. \tag{39}$$

From the equality of radiated and dissipated power:

$$\pi\rho\bar{r}=\frac{2R_r}{1+(\sin 2\beta\bar{r}\pi/2\beta\bar{r}\pi)}. \tag{40}$$

On the other hand[18]

$$\alpha\bar{r}\pi=\frac{\pi\rho\bar{r}}{2Z_0}=\frac{R_r}{Z_0[1+(\sin 2\beta\bar{r}\pi/2\beta\bar{r}\pi)]}, \tag{41}$$

where Z_0 is the characteristic impedance of equivalent transmission line. According to Libby[9]

$$Z_0=276\log_{10}(\bar{r}/d), \tag{42}$$

where $\bar{r}$ is the mean loop radius, and d the loop conductor diameter.

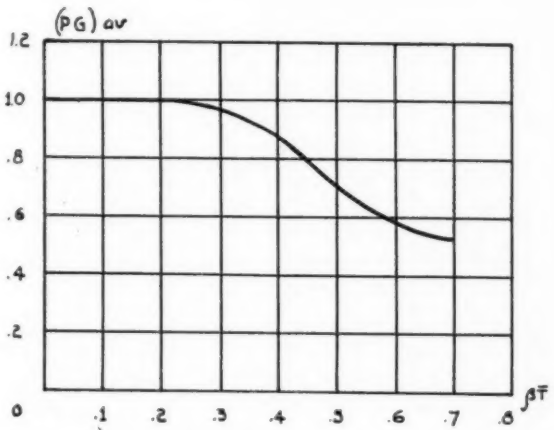


FIG. 6. Average power gain *versus* $\beta\bar{r}$ for $\alpha\bar{r}=0$.

ACKNOWLEDGMENTS

The author wishes to thank Mr. E. S. Kelsey for his help during the preparation of this paper, and Mr. J. C. Murray for computing aid.

APPENDIX I

Bessel Expansion of sin*t* and cos*t*

Since $\sin\theta\cos(\varphi-\bar{\varphi})=\frac{1}{2}(\sin\delta+\sin\epsilon)$, where $\delta=\theta+\varphi-\bar{\varphi}$, $\epsilon=\theta-\varphi+\bar{\varphi}$. Therefore,

$$\sin t=\sin\left[\frac{\beta\bar{r}}{2}\sin\delta\right]\cos\left[\frac{\beta\bar{r}}{2}\sin\epsilon\right]+\cos\left[\frac{\beta\bar{r}}{2}\sin\delta\right]\sin\left[\frac{\beta\bar{r}}{2}\sin\epsilon\right].$$

Using first terms of Bessel expansion[12] for sin and cos, $\sin t\leqq 4J_0J_1\sin\theta\cos(\varphi-\bar{\varphi})$, where $J_0\equiv J_0(\beta\bar{r}/2)$, $J_1\equiv J_1(\beta\bar{r}/2)$. Similarly,

$$\cos t=\cos[(\beta\bar{r}/2)\sin\delta]\cos[(\beta\bar{r}/2)\sin\epsilon]-\sin[(\beta\bar{r}/2)\sin\delta]\sin[(\beta\bar{r}/2)\sin\epsilon].$$

[18] S. A. Schelkunoff, reference 10, p. 196.

Using again first terms of Bessel expansion for sin and cos, $\cos t \leqq J_0^2$, since $J_1^2 \ll J_0^2$.

APPENDIX II

Definite Integrals Appearing in Formulas (6)

Using Dwight's "Tables of Integrals" and elementary trigonometric identities:

$$\int_0^\pi \cosh\gamma\hat{r}\cdot \sin\bar{\varphi}\cos(\varphi-\bar{\varphi})d\bar{\varphi} = \frac{1}{4+\gamma^2\hat{r}^2}\left[(1-\cos j\gamma\hat{r}\pi)\cos\varphi + \frac{2\sin j\gamma\hat{r}\pi}{j\gamma\hat{r}}\sin\varphi\right], \quad \text{(A2-1)}$$

$$\int_0^\pi \sinh\gamma\hat{r}\bar{\varphi}\sin\bar{\varphi}\cos(\varphi-\bar{\varphi})d\bar{\varphi} = \frac{j}{4+\gamma^2\hat{r}^2}\left[\sin j\gamma\hat{r}\pi\cos\varphi + \frac{2(\cos j\gamma\hat{r}\pi-1)}{j\gamma\hat{r}}\sin\varphi\right], \quad \text{(A2-2)}$$

$$\int_0^\pi \cosh\gamma\hat{r}\bar{\varphi}\sin\bar{\varphi}d\bar{\varphi} = \frac{1+\cos j\gamma\hat{r}\pi}{1+\gamma^2\hat{r}^2}, \quad \text{(A2-3)}$$

$$\int_0^\pi \sinh\gamma\hat{r}\bar{\varphi}\sin\bar{\varphi}d\bar{\varphi} = -j\frac{\sin j\gamma\hat{r}\pi}{1+\gamma^2\hat{r}^2}. \quad \text{(A2-4)}$$

Therefore, after reduction, formula (6) yields

$$N_x = I\hat{r}J_0\left[-jJ_1\sin\theta\frac{16\sin j\gamma\hat{r}\pi}{j\gamma\hat{r}(4+\gamma^2\hat{r}^2)}\sin\varphi\right]. \quad \text{(A2-5)}$$

Similarly

$$\int_0^{\pi} \cosh\gamma\hat{r}\varphi\cos\bar{\varphi}\cos(\varphi-\bar{\varphi})d\bar{\varphi} = \frac{1}{4+\gamma^2\hat{r}^2}\left[\frac{(2+\gamma^2\hat{r}^2)\sin j\gamma\hat{r}\pi}{j\gamma\hat{r}}\cos\varphi + (1-\cos j\gamma\hat{r}\pi)\sin\varphi\right], \quad \text{(A2-6)}$$

$$\int_0^{\pi} \sinh\gamma\hat{r}\bar{\varphi}\cos\bar{\varphi}\cos(\varphi-\bar{\varphi})d\bar{\varphi} = \frac{j}{4+\gamma^2\hat{r}^2}\left[\frac{(2+\gamma^2\hat{r}^2)(\cos j\gamma\hat{r}\pi-1)}{j\gamma\hat{r}}\cos\varphi + \sin j\gamma\hat{r}\pi\sin\varphi\right], \quad \text{(A2-7)}$$

$$\int_0^\pi \cosh\gamma\hat{r}\bar{\varphi}\cos\bar{\varphi}d\bar{\varphi} = \frac{j\gamma\hat{r}\sin j\gamma\hat{r}\pi}{1+\gamma^2\hat{r}^2}, \quad \text{(A2-8)}$$

$$\int_0^\pi \sinh\gamma\hat{r}\bar{\varphi}\cos\bar{\varphi}d\bar{\varphi} = j\frac{j\gamma\hat{r}(1+\cos j\gamma\hat{r}\pi)}{1+\gamma^2\hat{r}^2}. \quad \text{(A2-9)}$$

Therefore, after reduction, the second formula (6) yields

$$N_y = I\hat{r}J_0\left[jJ_1\sin\theta\frac{8(2+\gamma^2\hat{r}^2)\sin j\gamma\hat{r}\pi}{j\gamma\hat{r}(4+\gamma^2\hat{r}^2)}\cos\varphi - J_0\frac{2j\gamma\hat{r}\sin j\gamma\hat{r}\pi}{1+\gamma^2\hat{r}^2}\right]. \quad \text{(A2-10)}$$

APPENDIX III

Approximations to Formula (11)

Write

$$\gamma\hat{r} = \alpha\hat{r} + j\beta\hat{r} = j\beta\hat{r}[1-j(\alpha/\beta)].$$

Since $\alpha \ll \beta$

$$(\gamma\hat{r})^2 \leqq -(\beta\hat{r})^2\left[1-j\frac{2\alpha\hat{r}\beta\hat{r}}{(\beta\hat{r})^2}\right].$$

Therefore,

$$\frac{1}{1+\gamma^2\hat{r}^2} \leqq \frac{1}{1-(\beta\hat{r})^2} - j\frac{2\alpha\hat{r}\beta\hat{r}}{1-2(\beta\hat{r})^2} = a-jb,$$

$$\frac{1}{4+\gamma^2\hat{r}^2} \leqq \frac{1}{4-(\beta\hat{r})^2} - j\frac{2\alpha\hat{r}\beta\hat{r}}{16-8(\beta\hat{r})^2} = e-jf,$$

$$\frac{1}{\gamma^2\hat{r}^2(4+\gamma^2\hat{r}^2)} \leqq -\frac{1}{(\beta\hat{r})^4[16-8(\beta\hat{r})^2]} - j\frac{2\alpha\hat{r}\beta\hat{r}}{4(\beta\hat{r})^4} = c-jd.$$

APPENDIX IV

Integral in Formula for Radius of Equivalent Circular Horizontal Pattern

From formula (19)

$$|E_h|^2 = B^2[64(J_1/J_0)^2(d^2+c^2) + 16(J_1/J_0)(ad-bc)\cos\varphi + (a^2+b^2)\cos^2\varphi].$$

Therefore, after the elementary integration

$$\int_0^{2\pi} |E_h|^2 d\varphi = \pi B^2[128(J_1/J_0)^2(d^2+c^2) + (a^2+b^2)].$$

APPENDIX V

Integration of Formula (30)

Using Dwight's "Tables of Integrals," the integrals appearing in formula (30) are:

$$\int_0^\pi \sin^2 2\theta\sin\theta d\theta\int_0^{2\pi}\sin^2 2\varphi d\varphi = \frac{16}{15}\pi,$$

$$\int_0^\pi \sin 2\theta\sin\theta\cos\theta d\theta\int_0^{2\pi}\sin 2\varphi\sin\varphi d\varphi = 0,$$

$$\int_0^\pi \sin\theta\cos^2\theta d\theta\int_0^{2\pi}\sin^2\varphi d\varphi = \frac{2}{3}\pi,$$

$$\int_0^\pi \sin^3\theta d\theta\int_0^{2\pi} d\varphi = \frac{8}{3}\pi,$$

$$\int_0^\pi \sin^2\theta d\theta\int_0^{2\pi}\cos\varphi d\varphi = 0,$$

$$\int_0^\pi \sin\theta d\theta\int_0^{2\pi}\cos^2\varphi d\varphi = 2\pi.$$

High Speed Compression Tests on Copper

M. GREENFIELD, *North American Aviation, Los Angeles, California*

AND

E. T. HABIB, *David Taylor Model Basin, Carderock, Maryland*

(Received December 19, 1946)

Dynamic shortening of one-half inch long copper cylinders is achieved by striking them with a hardened steel projectile at high velocities. The average strain rate was about 1200 per second. Energy per unit volume absorbed by the copper is plotted against strain. A true stress-logarithmic strain curve is computed. This curve is compared with a similar curve derived from high speed tests on copper in tension.

INTRODUCTION

THIS work was undertaken originally to provide a dynamic calibration for copper cylinders. These cylinders are used in mechanical gauges which are designed to measure the pressure in the water originating from an underwater explosion. The copper is deformed dynamically in compression in such gauges. The experimental arrangement described below was designed so that the copper cylinders were subjected to the same kind of forces as they were when used in the gauges. However, it was thought that additional information could be obtained from the data which would be of wider interest.

These data have been put in the form of stress *versus* strain curves. Usually stress is defined as the force acting on the specimen divided by the original cross-sectional area of the test portion, i.e.,

$$\sigma_c = F/A_0$$

where A_0 is the original cross-sectional area. This stress, σ_c, is also called the engineering stress. The strain is frequently defined as the change in length per unit original length, i.e.,

$$\epsilon_c = \Delta l/l_0$$

where l_0 is the original length.

While some work has been done by Nadai and Manjoine on high speed tension tests of metals,[1] very little has been reported on high speed compression tests such as those described below. It was thought useful to compare the properties of copper in tension with those in compression when the forces are applied rapidly. The parameter chosen to describe the rapidity of the test is the time rate of strain $\dot{\epsilon} = d\epsilon_c/dt$.

For purposes of comparison of the properties of a metal in compression with those in tension, the choice of variables as defined above is poor. In tension tests the cross-sectional area of a specimen is diminished, making the true stress based on actual area greater than the engineering stress. In compression tests this area is increased, making the true stress less than the conventional stress. There is a similar difficulty in the comparison of strains for tension and compression tests. For these reasons it was decided to use the so-called true stress and logarithmic strain. For compression tests these quantities may be defined as follows:

$$\sigma = F/A$$

where σ is the true stress, F is the applied force, and A is the instantaneous cross-sectional area of the specimen.

$$\epsilon = \ln A_0/A,$$

where ϵ is the logarithmic strain, and A_0 is the original cross-sectional area of the specimen. The logarithm is in terms of the base e.

These quantities σ and ϵ are simply related to the conventional ones σ_c and ϵ_c, as follows:

$$\sigma = \sigma_c(1-\epsilon_c), \tag{1}$$

$$\epsilon = \ln 1/(1-\epsilon_c). \tag{2}$$

These relations are derived[2] by making the usual assumption for the plastic flow of metals

[1] M. Manjoine and A. Nadai, Proc. A.S.T.M. **40**, 822 (1940).

[2] A. Nadai, App. Phys. **8**, 205 (1937).

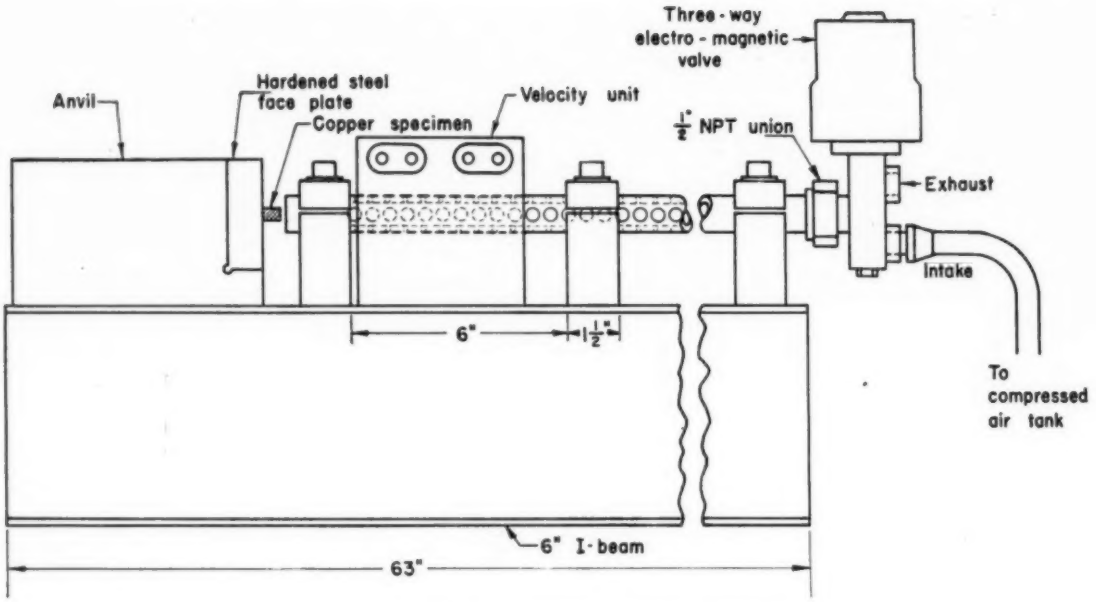


FIG. 1. Air gun for compressing copper cylinders.

that the volume of any element of metal remains unchanged.

APPARATUS

The test equipment consists of an air gun for blowing a hardened steel piston against the soft copper specimen and auxiliary equipment for measuring the energy of the piston (see Figs. 1 and 2). The gun barrel is a cylindrical tube of steel 55″ long with an outside diameter of $\frac{15}{16}$ inch and an inner diameter of $\frac{1}{2}$ inch. At the muzzle end of the barrel, spaced $\frac{1}{8}$ inch apart, one hundred eighty $\frac{1}{4}$″ holes are drilled in the barrel.

The gun barrel is laid in a V-shaped groove cut into a set of steel supports. Similarly grooved hold-down plates, bolted to the supports, keep the barrel in position. These supports are in turn bolted to a 6-inch I-beam, about 63 inches long. At the muzzle end of the barrel a steel anvil, with a hardened steel face plate, is bolted to the I-beam. The test specimen is placed against this steel plate. The face of the hardened steel plate is very nearly perpendicular to the line of travel of the piston. This insures that the faces of the copper specimen are almost parallel after deformation. The variation in readings at different points on the faces of the specimen is thus kept small.

The propulsive force for the piston is supplied by a compressed air system. Air from a tank at high pressure is led into a smaller tank through a regulating valve. This regulator enables the small tank to be filled to any pressure up to 400 lb. per square inch, although 100 lb./in.^{2} was the maximum used. A 3-way quick-acting solenoid valve controls the flow of air from the small reservoir to the gun barrel. The time of opening and closing of the valve is determined by a time-delay relay. The sequence of events is the following. A double-pole, double-throw switch is thrown. One side of the switch closes a circuit allowing 220-volt a.c. to flow through the solenoid; thus energized, the solenoid opens the intake valve and allows air to blow against the piston. The piston then moves down the barrel, strikes the specimen and rebounds. Simultaneously with the closure of the first side of the switch the second side closes a circuit in the time-delay unit. After a predetermined time delay a relay is thrown which opens the solenoid circuit. This closes the intake valve and, at the same time, opens the exhaust valve, allowing the rebounding piston to compress and force air out through the exhaust; in so doing, the piston brings itself to a halt.

By actual test, the shortest time for which the time-delay unit could be set to operate properly was about 0.01 second. This setting was used at

the fastest velocities and worked very well. In the other direction, the time-delay unit may be set to operate for as much as several seconds. The purpose in keeping the valve open for only a short time is to keep the excess of air flowing through the barrel to a minimum.

The test specimens are made from oxygen-free high conductivity copper rod. Each specimen is made so that the final length is 0.5000 inch ± 0.0005 inch and the diameter is 0.3255 inch ± 0.0005 inch. They are then annealed in an electrically heated furnace in an atmosphere of hydrogen at a temperature of 950° Fahrenheit. The temperature is maintained for $2\frac{1}{2}$ to 3 hours and the pellets allowed to cool in the furnace. The flow of hydrogen is maintained until the temperature is below 250° Fahrenheit.

The velocity of the piston is measured by noting the time interval between the breaking and making of a narrow beam of light which is incident upon a photo-cell (see Fig. 2). As the piston breaks the first light beam, a voltage change is induced in the photo-cell. This voltage is then amplified and used to trip a sweep generator. The sweep generator controls the horizontal movement of an electron beam across the face of a cathode-ray oscillograph. When the electron spot has completed about a third of its journey, the piston interrupts the second beam of light, producing a voltage impulse

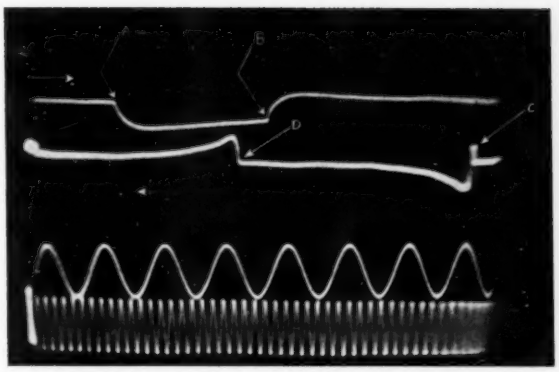


FIG. 3. A typical record obtained by the interruption of the light beam by the moving piston. At (A), the piston first enters the light beam; at (B), it is leaving the light beam. (C) and (D) represent a similar situation after the piston has rebounded. The lower oscillogram is the timing calibration for AB and CD.

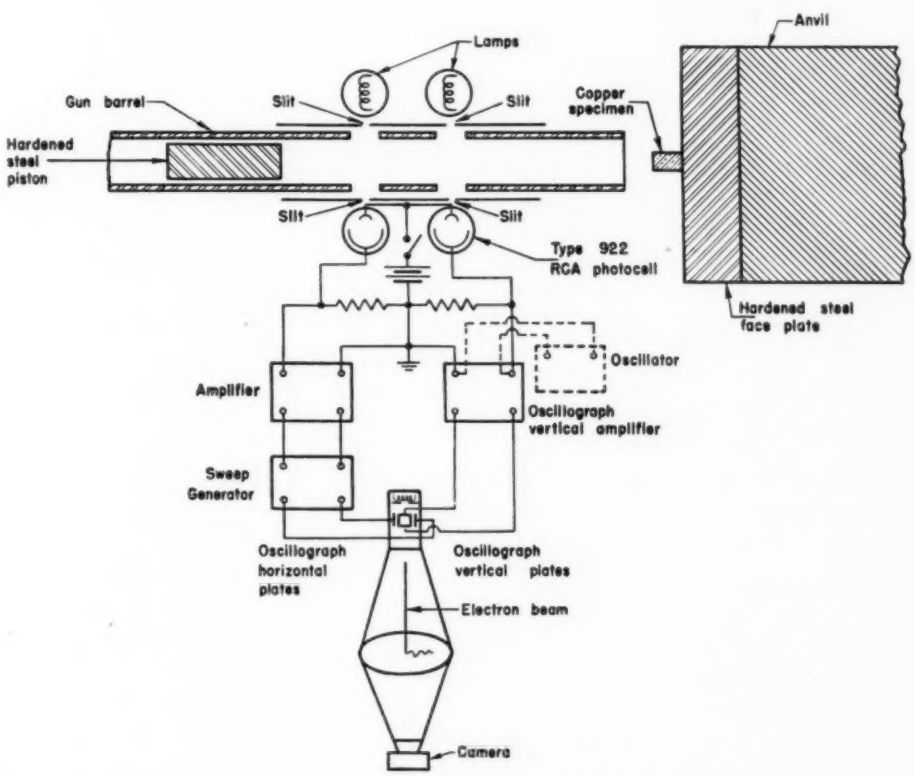


FIG. 2. Apparatus for measuring velocity of piston.

which is led into the vertical amplifier of the oscillograph. The spot then jumps up. When the piston has passed the light beam, the spot jumps back (see Fig. 3). After the piston strikes the copper, it rebounds; this velocity is recorded on the return trace of the electron beam. The distance between jumps is related to time by comparing it with a sine wave of known frequency. This sine wave is taken from an audio-frequency oscillator. To insure accurate time

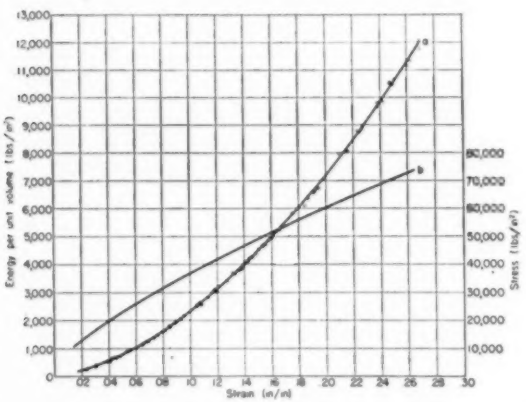


FIG. 4. Energy per unit volume-strain curve and stress-strain curve for copper cylinders in compression. Curve (a) shows dependence of energy absorbed per unit volume on the strain in compression in copper cylinders. Curve (b) is a stress-strain curve for copper derived from curve (a) by numerical differentiation. The rate of strain is approximately 1200 per second.

measurements, the output of the oscillator is beat against that of a 1000-cycle tuning-fork oscillator before the timing oscillogram is made. Knowing the piston length and its time to pass by the light beam, the velocity of the piston may be calculated. The energy of deformation is then computed from the relation K.E. $= \frac{1}{2}m(v_1^2 - v_2^2)$ where m is the mass of the piston, v_1 is its velocity before striking the specimen, and v_2 is its velocity of rebound. The oscilloscope tracings were photographed on 35-mm Super XX film in a camera equipped with an f:1.9 lens.

PRECISION OF MEASUREMENTS

Before being subjected to impact, each cylinder was measured with a micrometer. One reading was made at the center, four at intervals around the circumference. The average of these was taken as the length of the pellet before deformation. The averages of two sets of reading made in this way repeated to within 0.0001". After impact, the pellet was again measured at several points around the circumference and an average computed. (The center reading was discarded because it was consistently higher than the average.) The averages of two sets of readings in this case always agreed to within 0.0002". The maximum error possible from this measurement is then 0.0003". When the deformation is greater than 0.03", this error is less than 1 percent.

Measurements on the film are made with a traveling microscope. Measurements with this instrument can be made to a precision of 0.001 millimeter. The line thickness on the film was about 0.1 millimeter. The distance between jumps on the film was of the order of 5 millimeters. Measurements were made to the center of each line. Three or more readings were taken at each point and averaged. The averages could be repeated to within 0.01 millimeter. The maximum error possible in this measurement then is about 0.2 percent.

Measurements on the sine wave were made at the same portion of the film as the velocity record occurred. This was essential because the sweep was not exactly linear, especially at the edges of the scope face. However in the middle half of the oscillogram, the linearity was within 1 percent. The oscillator frequency was always made high enough so that about five cycles covered the same distance as the velocity record (see Fig. 3). Errors due to non-linearity of the sweep were thus reduced. A timing oscillogram was made before and after every few records. Velocities were computed using both timing records and an average taken. Velocities based on different timing oscillograms varied from the average by less than 0.5 percent.

The total maximum random error is thus about 2 percent.

STRAIN RATE ($\dot{\varepsilon}$)

Strain rate is defined as the time rate of change of the strain. Since the difference between the conventional strain and the logarithmic strain is small here, of order 10 percent, and since the dependence of stress on strain rate turns out to be a very slowly changing function (roughly multiplying the strain rate by ten means an increase of 4 or 5 percent in stress), it is suf-

ficiently accurate to use the time rate of change of the conventional strain for the strain rate. For the ½-inch long coppers used here, assuming the average velocity of the striking piston is one-half of the initial velocity with which it strikes the specimen, i.e., uniform deceleration, the average strain rate is numerically equal to the initial velocity expressed in inches/sec.

$$\text{Average } \dot{\epsilon} = \frac{\Delta l}{l_0} \cdot \frac{1}{T} = \frac{\Delta l}{T} \cdot \frac{1}{l_0}$$

$$= \frac{V_a}{l_0} = \frac{V_0}{2l_0} = V_0 \frac{\text{in./sec.}}{\text{in.}},$$

where Δl = final deformation, l_0 = original length = ½ inch, T = total time of deformation-seconds, v_0 = initial velocity-inches/second, and v_a = average velocity-inches/second.

The average strain rate is computed by assuming that the velocity with which the specimen contracts, i.e., the velocity of the striking piston, is a decreasing linear function of time. This assumption was checked by making two high speed photographs of the deformation of the copper cylinder. These showed that while the velocity of the striking piston is not linear, the true average velocity exceeded the assumed one by only about 15 percent. Because of the nature of the dependence of the stress on the strain rate, this method of computing the average strain rate is considered adequate.

DATA

Figure 4, curve (4a), is a plot of energy/unit volume in lb./in.² against conventional strain in inches/inch. It represents the basic data of the dynamic testing in compression. A smooth curve was drawn through the points in such a manner that as many points lay above the curve as below it. One copper specimen was used for each point and 118 points were obtained. Two hardened steel pistons were employed; one weighed 0.05473 lb., the other 0.1633 lb. Through their use, the average strain rate at the upper and lower ends of the curve was kept within about 50 percent of 1200 1/sec.

Curve (4b) which shows conventional stress *versus* conventional strain was derived from curve (4a) by numerical differentiation. In order to compare these compression data with data in tension, curve (4b) was revised by applying formulas (1) and (2).

Figure 5 shows the revised stress-strain curves. In all cases the ordinate is true stress and the abscissa is logarithmic strain. Curves (a) and (b) in this figure show the relation between stress and strain for copper in compression. All the specimens used in obtaining curves (a) and (b) came from the same lot. The strain rates that apply to these curves are approximately 1200 1/sec. for curve (a) and 1.5×10^{-3} 1/sec. for curve (b). The latter curve was obtained by means of an ordinary static compression test. Curve (a) was based on data from the dynamic test and was obtained by the procedure described above. The stress required to produce a given compressional strain dynamically was always greater than that required to produce the same strain statically. This difference in stress increases as the strain increases. Comparison of curves (a) and (b) shows that the percentage difference varies from 17 percent to 29 percent for the interval of strain from 0.10 to 0.30.

It was thought useful to compare these results with those obtained in high speed tension tests of copper at room temperatures. Some work of this kind has been done by Manjoine and Nadai.[1] Curves (c) and (d) in Fig. 5 show the stress-strain curves obtained by them for copper. Their specimens were made of commercially pure copper and were 0.200″ in diameter and 1″ in length. After being machined the specimens were annealed by heating them in dry regenerated gas to 500°C. After that they were cooled to room temperature in 12 hours.

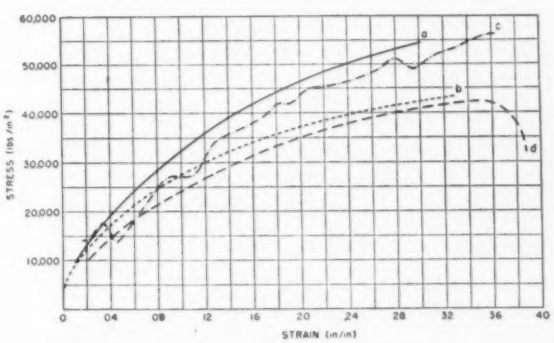


FIG. 5. Stress-strain curves for copper. The ordinate is true stress and the abscissa is logarithmic strain.

Curve (c) in Fig. 5 was obtained in a test of a

single specimen subjected to a rate of strain of 900 per second. Curve (d) in Fig. 5 was obtained by means of a static tension test. The difference between the two static curves (b) and (d) is probably due to the fact that the copper used by the writers was not quite the same as that used by Nadai and Manjoine. Comparison of curves (c) and (d) shows that the stress required to produce a given strain in tension dynamically generally exceeds that required to produce the same strain statically. Some of the oscillations in curve (c) are probably not real but may be due to the method of loading the specimen and measuring that load. A more valid curve may possibly be obtained by drawing a smooth average line through the oscillations. If one does this, then the percentage difference between curves (c) and (d) varies from approximately 19 percent to 25 percent for the interval of strain from 0.10 to 0.30. It is interesting to note that the percentage increase in stress required to produce a given strain in copper dynamically rather than statically is much the same for both compression and tension. This comparison should be made when the rates of strain are approximately the same.

$TM_{0,1}$ Mode in Circular Wave Guides with Two Coaxial Dielectrics

SIDNEY FRANKEL
Federal Telecommunication Laboratories, Inc., New York, New York
(Received February 3, 1947)

Field components for a transverse magnetic wave in a wave guide with two coaxial dielectrics are computed. A typical example is given to show the calculation of guide dimensions to reduce phase velocity to a preassigned value.

A UNIFORM circular wave guide may be conceived as a device that provides an interaction space between electrons and an electromagnetic field to transfer energy from one to the other. If electrons are injected axially in the guide and the field has an axial electric component, the important interaction is between the electrons and this axial component.

For unimpeded motion, the electrons must travel essentially in a vacuum and with a velocity necessarily less than that of light. On the other hand, the phase velocity of the electromagnetic field in an evacuated circular guide is always greater than the velocity of light. For effective interaction, this field must be slowed down, perhaps to the order of one-tenth or less of the velocity of light. This may be accomplished by surrounding the evacuated axial region of the guide with material of high dielectric constant. For simplicity, circular symmetry may be maintained.

An investigation of the propagation of an electromagnetic field with axial electric component in such a "loaded" guide has been made. Appropriate field equations, which take the necessary boundary conditions into account, are set up and solved. A relationship between the maximum value of axial electric component and transmitted power is given.

1. SOLUTION OF FIELD EQUATIONS

1.1 Differential Equations

Figure 1 shows a cross section of a wave guide. The section of the axial region is a circle of radius a, the constants of the medium being μ and ϵ_2, the permeability and dielectric constants, respectively. The boundary of the guide is a perfect conductor of radius b. In the annular region between these two radii is a material of constants μ and ϵ_1. It is assumed that $\epsilon_1 > \epsilon_2$ always.

Cylindrical coordinates r, ϕ, and x will be used with the x axis coinciding with the axis of the guide and being positive in the direction toward a sink of power. For a transverse magnetic mode, take $H_x = 0$, and, in accordance with the conventional $TM_{0,1}$ mode for uniform dielectrics, specify additionally that $E_\phi = H_r = 0$. If a

field satisfying these requirements exists in the guide under consideration, then the appropriate field equations in m.k.s. units for either dielectric are of the following type.

$$\left.\begin{aligned} -\partial H_\phi/\partial x &= j\omega\epsilon E_r, \\ (1/r)(\partial/\partial r)(rH_\phi) &= j\omega\epsilon E_x, \\ (1/r)(\partial E_x/\partial\phi) &= 0, \\ (\partial E_r/\partial x)-(\partial E_x/\partial r) &= -j\omega\mu H_\phi, \\ (1/r)(\partial E_r/\partial\phi) &= 0. \end{aligned}\right\} \quad (1)$$

The subscript has been omitted from ϵ to avoid restricting it to either medium.

According to (1), E_r and E_x, therefore also H_ϕ, are independent of ϕ. The useful equations from (1) are

$$\left.\begin{aligned} \partial H_\phi/\partial x &= -j\omega\epsilon E_r, \\ (1/r)(\partial/\partial r)(rH_\phi) &= j\omega\epsilon E_x, \\ (\partial E_r/\partial x)-(\partial E_x/\partial r) &= -j\omega\mu H_\phi. \end{aligned}\right\} \quad (2)$$

Elimination of E_r and E_x among these equations leads to the following equation in H_ϕ:

$$\frac{\partial^2 H_\phi}{\partial x^2}+\frac{\partial}{\partial r}\left[\frac{1}{r}\frac{\partial}{\partial r}(rH_\phi)\right]=-\omega^2\mu\epsilon H_\phi=-\frac{\omega^2}{c^2}H_\phi, \quad (3)$$

where c is the velocity of propagation in a medium of unbounded extent with constants μ and ϵ.

Assume

$$H_\phi=R(r)e^{-\gamma x}, \quad (4)$$

where γ is the propagation constant in the x direction and R is a function of r to be determined. Substitution of (4) in (3) yields

$$\frac{d}{dr}\left[\frac{1}{r}\frac{d}{dr}(rR)\right]+\alpha^2 R=0, \quad (5)$$

where

$$\alpha^2=\gamma^2+\omega^2/c^2. \quad (6)$$

Solutions of (5) are of the type

$$R=J_1(\alpha r),\ Y_1(\alpha r), \quad (7)$$

where J_1 and Y_1 are Bessel functions of the first and second kinds, respectively, and of the first order. If $Z_1(\alpha r)$ represents any linear combination of these functions, then

$$H_\phi=Z_1(\alpha r)e^{-\gamma x}. \quad (8)$$

For unattenuated transmission, γ is a pure imaginary; let $\gamma=j\beta$, with β real. Then

$$H_\phi=Z_1(\alpha r)e^{-j\beta x},\quad \alpha=[(\omega/c)^2-\beta^2]^{\frac{1}{2}}. \quad (9)$$

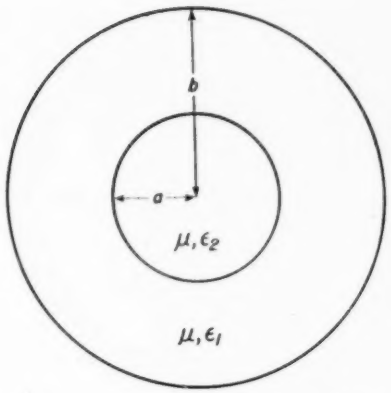


FIG. 1. Cross section of wave guide having two coaxial dielectrics.

From (2), we get immediately

$$\left.\begin{aligned} E_r &= (\beta/\omega\epsilon)Z_1(\alpha r)e^{-j\beta x}, \\ E_x &= (\alpha/j\omega\epsilon)Z_0(\alpha r)e^{-j\beta x}, \end{aligned}\right\} \quad (10)$$

where Z_0 is a linear combination of Bessel functions of the first and second kind, zeroth order, and we have made use of the relation[1]

$$\frac{d}{dr}[rZ_1(\alpha r)]=\alpha r Z_0(\alpha r). \quad (11)$$

J_0 and J_1 are finite and continuous everywhere in the complex plane, but Y_0 and Y_1 have infinite discontinuities at $r=0$. These latter functions are therefore inadmissable in the axial region of Fig. 1. The solutions therefore run as follows.

In medium 1,

$$\left.\begin{aligned} H_{\phi 1} &= [A_1 J_1(\alpha_1 r)+B_1 Y_1(\alpha_1 r)]e^{-j\beta_1 x}, \\ E_{r1} &= \frac{\beta_1}{\omega\epsilon_1}[A_1 J_1(\alpha_1 r)+B_1 Y_1(\alpha_1 r)]e^{-j\beta_1 x}, \\ E_{x1} &= \frac{\alpha_1}{j\omega\epsilon_1}[A_1 J_0(\alpha_1 r)+B_1 Y_0(\alpha_1 r)]e^{-j\beta_1 x}. \end{aligned}\right\}$$

$$a\leqq r\leqq b. \quad (12a)$$

In medium 2,

$$\left.\begin{aligned} H_{\phi 2} &= A_2 J_1(\alpha_2 r)e^{-j\beta_2 x}, \\ E_{r2} &= (\beta_2/\omega\epsilon_2)A_2 J_1(\alpha_2 r)e^{-j\beta_2 x}, \\ E_{x2} &= (\alpha_2/j\omega\epsilon_2)A_2 J_0(\alpha_2 r)e^{-j\beta_2 x}, \end{aligned}\right\} 0\leqq r\leqq a. \quad (12b)$$

[1] Jahnke and Emde, *Tables of Functions* (B. G. Teubner, Leipzig, 1938), third edition, p. 145.

A_1, B_1, and A_2 are constants to be determined from the boundary and power conditions, α_1 and α_2 are the propagation constants, and β_1 and β_2 are the phase constants, in media 1 and 2, respectively.

1.2 Boundary Conditions

Matching of tangential components of the electric field at the dielectric interface requires that

$$E_{z1}=E_{z2} \quad \text{at} \quad r=a$$

for all values of x; i.e.,

$$\frac{\alpha_1}{j\omega\epsilon_1}[A_1J_0(\alpha_1a)+B_1Y_0(\alpha_1a)]e^{-j\beta_1z} = \frac{\alpha_2}{j\omega\epsilon_2}A_2J_0(\alpha_2a)e^{-j\beta_2z}. \quad (13)$$

This can be satisfied for all values of x if, and only if, $\beta_1=\beta_2=\beta$. The phase constant (hence the velocity of propagation) is the same in both media, as might have been expected.

$$\alpha_1^2-\omega^2/c_1^2=-\beta^2=\alpha_2^2-\omega^2/c_2^2. \quad (14)$$

Assume the phase velocity V_p to be intermediate between c_1 and c_2

$$c_1<V_p<c_2, \quad V_p=\omega/\beta. \quad (15)$$

This leads to the inequality

$$\alpha_2^2<0<\alpha_1^2, \quad (16)$$

whence α_1 is real while α_2 is a pure imaginary. The field in medium 2 therefore behaves according to Bessel functions of the first kind with pure imaginary argument. Curves for $J_0(jz)$ and $-jJ_1(jz)$, frequently designated as $I_0(z)$ and $I_1(z)$, respectively, have been published.[2] The function $I_0(z)$ is particularly interesting since it shows the radial variation of the axial electric field. This field is fairly uniform for sufficiently small z. For $z<0.6$, the variation is less than 10 percent; for $z<2$, the variation is less than about 2:1.

Next, for the tangential component of electric field to vanish at the conductor surface, it is required that $E_{z1}=0$ at $r=b$, i.e.,

$$A_1J_0(\alpha_1b)+B_1Y_0(\alpha_1b)=0. \quad (17)$$

Matching the normal components of dielectric flux at the dielectric interface requires that

$$\epsilon_1E_{r1}=\epsilon_2E_{r2} \quad \text{at} \quad r=a; \quad \text{i.e.,}$$
$$A_1J_1(\alpha_1a)+B_1Y_1(\alpha_1a)=A_2J_1(\alpha_2a). \quad (18)$$

Matching the tangential magnetic field at the dielectric interface yields no further information. Collecting and rearranging (13), (17), and (18), we have the set of linear homogeneous equations in A_1, B_1, and A_2:

$$\left.\begin{aligned} \frac{\alpha_1}{\epsilon_1}J_0(\alpha_1a)A_1+\frac{\alpha_1}{\epsilon_1}Y_0(\alpha_1a)B_1 - \frac{\alpha_2}{\epsilon_2}J_0(\alpha_2a)A_2&=0,\\ J_0(\alpha_1b)A_1+Y_0(\alpha_1b)B_1&=0,\\ J_1(\alpha_1a)A_1+Y_1(\alpha_1a)B_1-J_1(\alpha_2a)A_2&=0. \end{aligned}\right\} \quad (19)$$

A necessary and sufficient condition that these equations be consistent is that the determinant of the coefficients vanish:

$$\begin{vmatrix} \frac{\alpha_1}{\epsilon_1}J_0(\alpha_1a) & \frac{\alpha_1}{\epsilon_1}Y_0(\alpha_1a) & -\frac{\alpha_2}{\epsilon_2}J_0(\alpha_2a) \\ J_0(\alpha_1b) & Y_0(\alpha_1b) & 0 \\ J_1(\alpha_1a) & Y_1(\alpha_1a) & -J_1(\alpha_2a) \end{vmatrix}=0. \quad (20)$$

If β is specified in advance, then α_1 and α_2 are known from (14), and (20) specifies a necessary and sufficient relationship between a and b. The manner of using this information will be illustrated by a later example.

With (20) satisfied, any two of the three relations given in (19) may be used to solve for the ratios of the constants A_1, B_1, and A_2. By using the first and the third,

$$A_1=K_AA_2, \quad B_1=K_BA_2, \quad (21)$$

where

$$\left.\begin{aligned} K_A&=-\frac{\pi\epsilon_1a}{2}\begin{vmatrix} \frac{\alpha_2}{\epsilon_2}J_0(\alpha_2a), & \frac{\alpha_1}{\epsilon_1}Y_0(\alpha_1a) \\ J_1(\alpha_2a), & Y_1(\alpha_1a) \end{vmatrix}\\ K_B&=-\frac{\pi\epsilon_1a}{2}\begin{vmatrix} \frac{\alpha_1}{\epsilon_1}J_0(\alpha_1a), & \frac{\alpha_2}{\epsilon_2}J_0(\alpha_2a) \\ J_1(\alpha_1a), & J_1(\alpha_2a) \end{vmatrix} \end{aligned}\right\} \quad (22)$$

and we have made use of the standard identity

$$Y_0(s)J_1(s)-Y_1(s)J_0(s)=2/\pi s.$$

[2] Page 224 of reference 1.

The field components are then given as follows:

Medium 1: $a \leqq r \leqq b$.

$$\left.\begin{aligned} E_{x1} &= A_2 \frac{\alpha_1}{j\omega\epsilon_1}[K_A J_0(\alpha_1 r) + K_B Y_0(\alpha_1 r)]e^{-j\beta x}, \\ E_{r1} &= A_2 \frac{\beta}{\omega\epsilon_1}[K_A J_1(\alpha_1 r) + K_B Y_1(\alpha_1 r)]e^{-j\beta x}, \\ H_{\phi 1} &= A_2[K_A J_1(\alpha_1 r) + K_B Y_1(\alpha_1 r)]e^{-j\beta x}, \end{aligned}\right\} \tag{23a}$$

Medium 2: $0 \leqq r \leqq a$.

$$\left.\begin{aligned} E_{x2} &= A_2 \frac{\alpha_2}{j\omega\epsilon_2} J_0(\alpha_2 r)e^{-j\beta x}, \\ E_{r2} &= A_2 \frac{\beta}{\omega\epsilon_2} J_1(\alpha_2 r)e^{-j\beta x}, \\ H_{\phi 2} &= A_2 J_1(\alpha_2 r)e^{-j\beta x}. \end{aligned}\right\} \tag{23b}$$

In (23), the constant A_2 remains to be determined.

1.3 Axial Field as a Function of Transmitted Power

The constant A_2 may be evaluated in terms of the peak transmitted power P_M or twice the average power P. For this purpose, the complex Poynting vector is integrated over a convenient cross section of the guide. This reduces to

$$P_M = \int_0^b \int_0^{2\pi} E_r H_\phi^* \cdot r\, d\phi\, dr,$$

where H_ϕ^* is the complex conjugate of H_ϕ. By (22), K_A and K_B both contain the factor α_2 and are therefore pure imaginaries. If E_r and H_ϕ are specified to have zero phase angle at $x=0$, then A_2 is also a pure imaginary, and j appears only in the exponential factor. The integration then runs

$$P_M = 2\pi A_2^2 \frac{\beta}{\omega\epsilon_1} F(a,b;\epsilon_1,\epsilon_2),$$

where

$$F = \frac{\epsilon_1}{\epsilon_2}\int_0^a r J_1^2(\alpha_2 r)\,dr + \int_a^b r[K_A J_1(\alpha_1 r) + K_B Y_1(\alpha_1 r)]^2\,dr \tag{24}$$

or

$$A_2 = ((P_M/2\pi)(\omega\epsilon_1/\beta F))^{\frac{1}{2}}. \tag{25}$$

At the center of the tube, the maximum value of the axial field E_{x2} is

$$E_{x0} = \left| A_2 \frac{\alpha_2}{\omega\epsilon_2} \right| = \left| \frac{\alpha_2}{\omega\epsilon_2} \left(\frac{P_M}{2\pi} \frac{\omega\epsilon_1}{\beta F} \right)^{\frac{1}{2}} \right|. \tag{26}$$

1.4 Further Discussion of the Solution

The determinant (20) may be considered to yield a solution for b when α_1 and α_2 are known. After defining a set of functions

$$Q_1(s) = \frac{J_0(s)}{Y_0(s)}, \quad Q_2(s) = -j\frac{J_1(js)}{J_0(js)}, \quad Q_3(s) = \frac{J_0(s)}{J_1(s)}, \quad Q_4(s) = \frac{Y_1(s)}{J_1(s)}, \tag{27}$$

we can reduce (20) to the form

$$Q_1(\alpha_1 b)=\frac{1-j\frac{\alpha_1}{\alpha_2}\frac{\epsilon_2}{\epsilon_1}Q_2(\alpha_2 a)Q_3(\alpha_1 a)}{Q_4(\alpha_1 a)-j\frac{\alpha_1}{\alpha_2}\frac{\epsilon_2}{\epsilon_1}Q_2(\alpha_2 a)\frac{Q_3(\alpha_1 a)}{Q_1(\alpha_1 a)}}. \tag{28}$$

All of these Q functions, with the exception of $Q_2(s)$ have been plotted.[3] $Q_2(s)$ is plotted in Fig. 2 of this paper. The use of these curves simplifies the calculation of b.

The chief remaining problem is an evaluation of the function $F(a, b; \epsilon_1, \epsilon_2)$ in (24). The function is integrable in known Bessel functions. By reference to a standard treatise on Bessel functions,[4] it can be shown that

$$F=\frac{1}{\alpha_2^2}F_1(\alpha_2 a)+\frac{K_A^2}{\alpha_1^2}[F_1(\alpha_1 b)-F_1(\alpha_1 a)]+\frac{2K_AK_B}{\alpha_1^2}[F_3(\alpha_1 b)-F_3(\alpha_1 a)]+\frac{K_B^2}{\alpha_1^2}[F_2(\alpha_1 b)-F_2(\alpha_1 a)], \tag{29}$$

where

$$\left.\begin{aligned} F_1(s)&=\frac{s^2}{2}\{[J_1(s)]^2-J_0(s)J_2(s)\},\\ F_2(s)-F_2(u)&=\frac{s^2}{2}\{[Y_1(s)]^2-Y_0(s)Y_2(s)\}-\frac{u^2}{2}\{[Y_1(u)]^2-Y_0(u)Y_2(u)\},\\ F_3(s)&=\tfrac{1}{2}\{s^2[Y_1(s)J_1(s)+Y_2(s)J_2(s)]-s[Y_1(s)J_2(s)+Y_2(s)J_1(s)]\}. \end{aligned}\right\} \tag{30}$$

Future work may indicate whether or not these results can be simplified.

2. TYPICAL EXAMPLE

Suppose a $TM_{0,1}$ wave is to be set up at $\omega=3\cdot10^{10}$ per second traveling at one-tenth the velocity of light in vacuum, so that $V_p=3\cdot10^7$ meters per second, whence $\beta=(10)^3$ per meter. Assume the dielectric material is such that

$\mu=4\pi10^{-7}$ henry per meter,
$\epsilon_1=(10^{-7}/18\pi)$ farad per meter
(relative dielectric constant = 200),
$\epsilon_2=(10^{-9}/36\pi)$ farad per meter.

Then $\alpha_1=10^3$, $\alpha_2=j10^3$ by (14). Next, let us try using a radius $a=2$ millimeters $=2\cdot10^{-3}$ meter. Then

$$\alpha_1 a=2,$$
$$\alpha_2 a=j2,$$
$$\alpha_1/\epsilon_1=5.55\cdot10^{11},$$
$$\alpha_2/\epsilon_2=j1.11\cdot10^{14}.$$

From Fig. 2, $Q_2(j2)=0.70$, and from reference 1 and (27)

$$Q_1(2)=0.45,$$
$$Q_3(2)=0.39,$$
$$Q_4(2)=-0.19.$$

Then by (28), $Q_1(\alpha_1 b)=-5.2$ whence

$$\alpha_1 b=0.73,\ 3.78,\ 6.89,\ \cdots$$

and

$$b=0.7,\ 3.8,\ 6.9,\ \cdots,\ \text{millimeters}.$$

We take $b=3.8$ millimeters as the smallest available value of $b>a$. A more careful calculation using tables rather than curves shows that $b=3.6$ millimeters.

Next, compute K_A and K_B using (22).

$$K_A=j153.3,$$
$$K_B=j812.5.$$

The expressions for the field components (23)

[3] Pages 200–203 of reference 1.
[4] G. N. Watson, *A Treatise on the Theory of Bessel Functions* (Macmillan Company, New York, 1944), second edition.

then become

$$E_{x1}=A_2 2842[J_0(\alpha_1 r)+5.30\,Y_0(\alpha_1 r)]e^{-j\beta z},$$
$$E_{r1}=jA_2 2892[J_1(\alpha_1 r)+5.30\,Y_1(\alpha_1 r)]e^{-j\beta z},$$
$$H_{\phi 1}=jA_2 153.3[J_1(\alpha_1 r)+5.30\,Y_1(\alpha_1 r)]e^{-j\beta z},$$
$$E_{x2}=A_2 3703 J_0(\alpha_2 r)e^{-j\beta z},$$
$$E_{r2}=A_2 3774 J_1(\alpha_2 r)e^{-j\beta z},$$
$$H_{\phi 2}=A_2 J_1(\alpha_2 r)e^{-j\beta z}.$$

These quantities are plotted as functions of r in Figs. 3 and 4. Here it is plain that the electric vectors have the same orders of magnitude in both dielectrics, while the magnetic vector is practically negligible in the central dielectric. This fact, combined with the small cross-sectional area of the inner dielectric, results in most of the transmitted power being propagated in the outer dielectric.

The factor F in (24) may be evaluated by numerical integration, whence it turns out that $F=-0.00037-0.3475=-0.3479$, the smaller term being the contribution of the inner dielectric. Since by (26) the transmitted power is related linearly to F, it can be seen that only about 0.1 percent of the power is propagated in the central region.

By substitution in (26) we now have immediately $E_{z0}=5.77(P_M)^{\frac{1}{2}}$, volts per centimeter, at the axis of the guide.

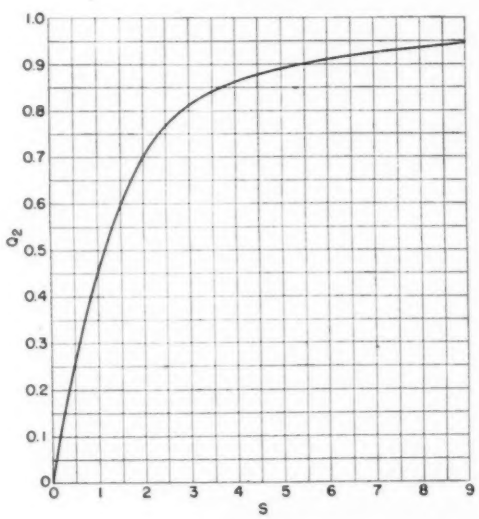


FIG. 2. A useful ratio of certain Bessel functions of imaginary argument. $Q_2=-j[J_1(js)/J_0(js)]$.

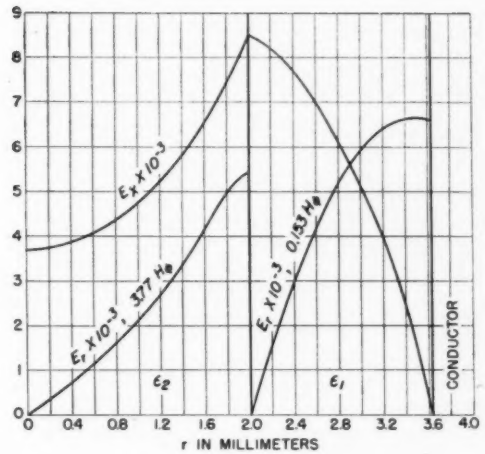


FIG. 3. Distribution of field components in a typical example.

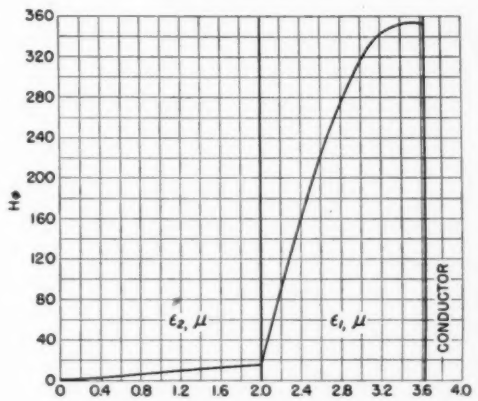


FIG. 4. Distribution of magnetic intensity in a typical example.

Graphical Methods for Evaluating Fourier Integrals

W. J. CUNNINGHAM
*Cruft Laboratory, Harvard University, Cambridge, Massachusetts**
(Received March 12, 1947)

The Fourier transform method of determining the response of a linear system to an arbitrary input signal often has its practical usefulness impaired because of difficulties in evaluating the necessary integrals. One possibility of overcoming these difficulties lies in the application of graphical methods to the transformations. Three such graphical procedures are described, all based upon fundamental properties of the transforms. Each method involves an analysis of the curves of the function to be transformed as a sum of curves of simpler functions whose transforms are known. The methods are useful in cases where the problem is too complicated for a simple analytic solution, or where part of the necessary data is available only in the form of a curve obtained, say, from experimental measurements of transmission characteristics or wave shapes. The accuracy of the methods is restricted only by that of the graphical plotting and curve fitting. If only approximate results are required, they may be obtained relatively quickly by these methods.

I. INTRODUCTION

ONE of the standard mathematical methods for determining the response of a linear transmission system to an input disturbance of arbitrary shape is based upon the Fourier integral.[1,2] In the application of this method it is necessary first to express the input function of time as a corresponding function of frequency by means of the appropriate Fourier integral. The resulting frequency function, or transform, is multiplied by the steady-state transmission characteristics of the system to give the frequency function at the output of the system. The time function, or inverse transform, at the output is obtained from the output frequency function by means of the inverse form of the Fourier integral. The integrals needed in these transformations may be written in a number of ways, a convenient form being as follows:

$$g(f)=\int_{-\infty}^{\infty} e(t)\exp(-j\omega t)dt, \qquad (1)$$

$$e(t)=\int_{-\infty}^{\infty} g(f)\exp(+j\omega t)df, \qquad (2)$$

$$\omega=2\pi f, \quad j=+(-1)^{\frac{1}{2}},$$

where $g(f)$ and $e(t)$ are corresponding transforms of frequency and of time. In Fourier integrals of this form, the entire frequency and time scales, both positive and negative values, must be considered.

Several difficulties may arise when an effort is made to apply this method to the solution of practical problems. Often the input signal is known in the form of a curve, but a mathematical expression for it is not available. The same thing may be true (perhaps more often so) for the transmission characteristics of the system. Even if these mathematical expressions were known, it might not be possible to carry out the integrations needed in the evaluation of the Fourier transforms. A number of standard forms of these integrals are available,[3] but if the unknown system does not fit one of these transforms, solution may be very difficult.

II. GRAPHICAL METHODS

A possible way of obviating these difficulties is to resort to a graphical analysis which will replace the integrations normally required. Furthermore, such graphical methods allow the use of data known only as a curve found from experimental measurements. If extremely accurate results are not required, approximations may be introduced in such a way that the graphical method may yield a useful result more quickly than other rigorous methods. The graphical method is particularly useful when such

* Now at Department of Electrical Engineering, Yale University, New Haven, Connecticut.

[1] E. A. Guillemin, *Communication Networks* (John Wiley and Sons, 1935), Vol. II, Chap. XI.

[2] W. L. Sullivan, "Analysis of systems with known transmission frequency characteristics by Fourier integrals," Elec. Eng. 61, 248 (1942).

[3] G. A. Campbell and R. M. Foster, "Fourier integrals for practical applications," Bell Telephone Monograph, B-584, (1931).

approximations are being made, since it is relatively easy to visualize their effect upon the final results. When the system is complicated, the graphical solution likewise becomes complicated, as does any other means of transient analysis. At least two graphical methods for determining the response of a system to an arbitrary input signal have already been described.[4,5] However, the procedures of the present discussion differ from these two and may have certain advantages in their application.

These methods are based upon fundamental properties of the Fourier integrals.[3] Perhaps the most important property is the superposition principle for linear systems. This principle states that if a function is made up of any linear combination of other functions, the Fourier transform of the first function is equal to the same linear combination of the Fourier transforms of the other functions. Therefore, it is possible to collect as geometric curves a number of mathematically simple functions and their transforms. A function in the form of a complicated curve may be broken into some linear combination of these simple curves. The transform of the complicated curve is then the same linear combination of the transforms of the simple curves. The whole process may be carried out by graphical methods.

In physical problems the input function of time and the output function of time are generally both real. Their corresponding transforms are usually complex functions, each having both a real and an imaginary part. The real parts of these complex functions have even symmetry; their inverse transforms are also real and have even symmetry. The imaginary parts of the complex functions have odd symmetry; their inverse transforms are real and have odd symmetry. Thus, the inverse transform of a complex function is real, and consists of the sum of an even function and an odd function.

An important type of problem is that in which the input to the system is zero until zero time, when the input function begins to have a finite value. If the physically realizable system is passive, its output must also be zero for times less than zero. This fact, plus the discussion of the preceding paragraph, leads to an important conclusion. The real time function at the output of the system must be zero for negative times and will generally have a non-zero value for positive times. Therefore, the even and odd real functions which compose it must have identical geometric shape at positive times and be mirror images of one another at negative times.[6] If this property is utilized, it is often necessary to consider only one of these component functions, and not both simultaneously.

The solution of a typical problem might proceed as shown in Fig. 1, where both positive and negative scales of time and frequency are shown. The input function of time, Fig. 1(a), is applied at zero time and is assumed to be sufficiently simple mathematically so that its transform, Fig. 1(b), may be found as a complex function of frequency from Eq. (1). If this is not true, the problem is more involved, but still may be solved by the methods to be described. It will also be assumed that the steady-state transmission characteristics of the system, Fig. 1(c), are known as a complex function of frequency. These characteristics may be found by the usual steady-state analysis methods, or in the case of actual systems, may be determined experimentally. It is necessary now to find the product, Fig. 1(d), at each frequency, of the complex functions of frequency representing the

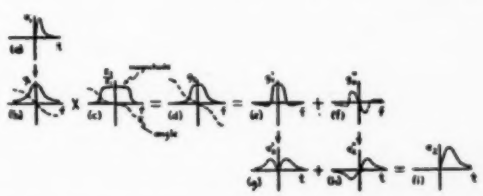


FIG. 1. Outline of solution of typical problem by Fourier integral method. Necessary mathematical operations are shown, with the arrow indicating a (direct or inverse) Fourier transformation. (a) Input signal to system (real). (b) Fourier transform of (a) (complex). (c) Steady-state transmission characteristics of system (complex). (d) Product of (b) and (c) (complex). (e) Real part of (d) (even symmetry). (f) Imaginary part of (d) (odd symmetry). (g) Inverse transform of (e) (real and even). (h) Inverse transform of (f) (real and odd). (i) Output signal from system (real).

[4] A. V. Bedford and G. L. Fredendall, "Transient response of multistage video frequency amplifiers," Proc. I.R.E. **27**, 277 (1939).

[5] H. A. Wheeler, "The interpretation of amplitude and phase distortion in terms of paired echoes," Proc. I.R.E. **27**, 384 (1939).

[6] E. C. Cherry, "Pulse response, a new approach to a.c. electric network theory and measurement," J. Inst. Elec. Eng. **92**, III, 183 (1945).

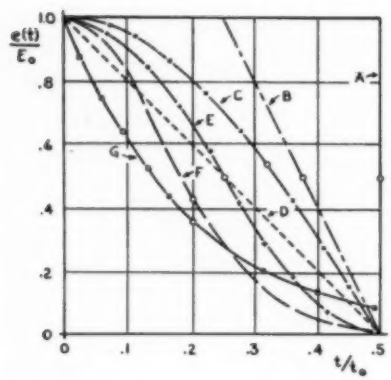


FIG. 2. Real even functions of time; only positive half of each is shown. The maximum value of the function is E_0; its total duration is approximately t_0. Circles indicate the nominal duration, t_n. Corresponding transforms are shown in Fig. 3. A. Rectangle. B. Trapezoid. C. Half cosine. D. Triangle. E. Cosine squared. F. Gaussian. G. Exponential.

input signal and the transmission characteristics of the system. If these two functions are available in algebraic form (magnitude and angle are most convenient), they may be multiplied directly. If one or both of them is available only as a curve, the multiplication of ordinates must be done on a point by point basis. The product obtained in this way will generally be a complex function of frequency and should be expressed as a real part, Fig. 1(e), and an imaginary part, Fig. 1(f). The inverse transform of each part, Fig. 1(g) and (h), is the contribution of that part to the real function of time at the output of the system. As previously pointed out, each contribution must be the same for positive values of time. Because of their even and odd symmetry, their sum at negative times is zero. In order to obtain the output signal, Fig. 1(i), the two contributions may be added, or more simply, only one contribution (either (g) or (h)) need be considered at positive times only, and its value doubled.

The last step of transforming the function of frequency at the output of the system into the corresponding function of time is usually the step producing the greatest mathematical difficulty. It is for this purpose that the methods described here will be found useful.

III. CORRESPONDING FUNCTIONS

A number of corresponding simple functions of time and of frequency are given in Fig. 2 and Fig. 3. A collection of curves of this sort is useful, either in the direct graphical solution of more complicated wave shapes, or in providing the investigator with a qualitative picture of the relation between typical functions. All of these functions of time and frequency are real and, therefore, must have even symmetry. Only one half of each curve is plotted, for the sake of convenience, but it must be remembered that each of these curves is symmetrical about the vertical axis and has the same values for negative as for positive abscissas. The curves are obtained by assuming relatively simple mathematical functions of time and finding the corresponding functions of frequency with the aid of Eq. (1). For convenience, the curves are plotted with dimensionless ratios as coordinates, and the time functions (Fig. 2) are adjusted so that the approximate total duration of each is the time t_0. Also indicated on the curves, with small circles, are the points corresponding to the nominal duration, t_n, and the nominal cut-off frequency, f_n, for each curve.[5] The nominal duration and nominal cut-off may be defined for real even functions in the following way,

$$\left.\begin{aligned} t_n &= 1/e(0)\int_0^\infty e(t)dt, \\ f_n &= 1/g(0)\int_0^\infty g(f)df, \end{aligned}\right\} \qquad (3)$$

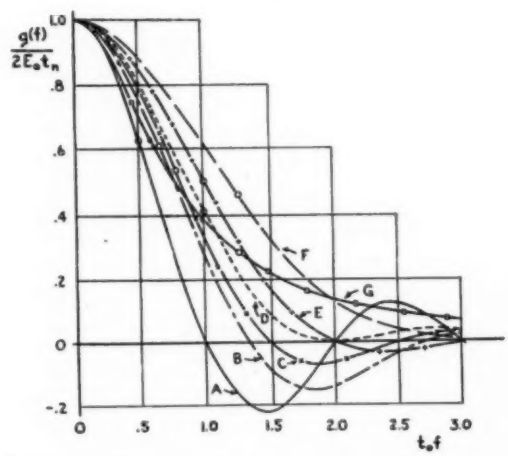


FIG. 3. Real even functions of frequency; only positive half of each curve is shown. Circles indicate the nominal cut-off frequency, f_n. The corresponding time functions have a maximum value, E_0, a nominal duration, t_n, and an approximate total duration, t_0, and are shown in Fig. 2. The frequency functions correspond to time functions of shape: A. Rectangle. B. Trapezoid. C. Half cosine. D. Triangle. E. Cosine squared. F. Gaussian. G. Exponential.

where $e(0)$ is $e(t)$ evaluated at zero time, and $g(0)$ is $g(f)$ evaluated at zero frequency, and both $e(0)$ and $g(0)$ are finite. In geometrical terms, the nominal duration, t_n, is the base of a rectangle which has an altitude, $e(0)$, and an area equal to the total area under the curve of $e(t)$, taken for positive times only. A similar relation applies to f_n. For zero values of the arguments, the quantities are

$$e(0)=\int_{-\infty}^{\infty} g(f)df=2f_n g(0),$$

$$g(0)=\int_{-\infty}^{\infty} e(t)dt=2t_n e(0),$$

and their product gives the fundamental relation,

$$4f_n t_n=1.$$

If this relation is observed and the parameters, E_0, t_0, and t_n, are chosen properly, the scales of Figs. 2 and 3 may be made to fit any desired conditions or systems of units. The maximum value of the time function, occurring here at zero time, is E_0.

One further important property of the transforms may be found from inspection of the Fourier integrals. If the arguments of the time function and the frequency function are interchanged, the functions will still correspond provided the sign of one argument is changed. With real even functions, such as Fig. 1(e) and (g), this sign change produces no visible change in the functions. Therefore, the real even functions of Fig. 2 may be taken as functions of frequency, if desired, whereupon the real even functions of Fig. 3 become functions of time. It is sufficient merely to interchange the t and f symbols of these curves.

Another group of corresponding functions which is of interest is shown in Fig. 4. Here, the functions of time, Fig. 4(a) and (c), are cosine functions in which an integral number, n, of quarter cycles occurs within the arbitrary time, t_a. Since the cosines are even functions the total duration of the functions of time is $2t_a$. Again, the frequency functions, Fig. 4(b) and (d), corresponding to the assumed time functions, may be found with the help of Eq. (1). The form of the result is slightly different depending upon whether n is chosen odd or even. Two families of curves are plotted in Fig. 4 for this reason. These frequency functions are the frequency spectra of a.c. pulses containing a total of n half-cycles, and occurring within the time $2t_a$. As the

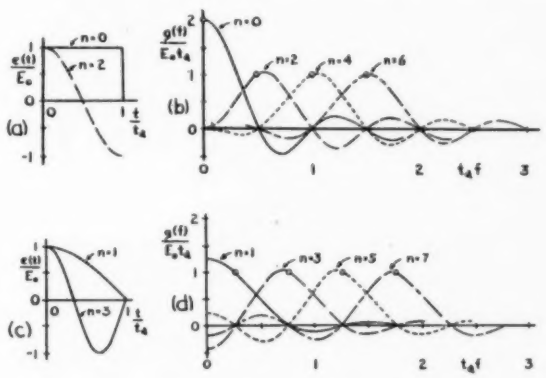


FIG. 4. Corresponding real even functions of time and frequency; only positive half of each curve is shown. Time functions have n quarter cycles in the interval, t_a. The parameter, n, is even at (a) and odd at (c). Corresponding frequency functions are shown at (b) and (d), respectively. Circles indicate points where only a single curve is not zero.

value of n is made larger, the apparent frequency of the time function is increased, and the main hump of the corresponding frequency spectrum is moved to a point higher on the frequency scale. All of these functions are real and even, although only the positive half of each curve is shown, and once more it is possible to interchange time and frequency if desired.

An important feature of these frequency functions, which is used to simplify later analyses, becomes apparent after a study of Fig. 4. If n is chosen either odd or even, there is a certain frequency at which the frequency function for a single value of n is not zero, while the functions for all other values of n are zero. The function which is not zero at this particular frequency has a value which is a constant, regardless of n. In mathematical terms, if n is an integer, either odd or even as the case may be, and m is an integer odd or even to match n, then

$$\left.\begin{array}{lll} \text{at} & f=m/(4t_a), & g(f)=0, \quad \text{if} \quad m\neq n \\ \text{at} & f=n/(4t_a), & g(f)=E_0t_a, \quad \text{if} \quad n\neq 0 \\ \text{at} & f=0, & g(f)=2E_0t_a, \quad \text{if} \quad n=0. \end{array}\right\} \quad (4)$$

The points where $m=n$ are shown by circles on Fig. 4(b) and (d). In one of the uses of these functions it is necessary to sum a number of

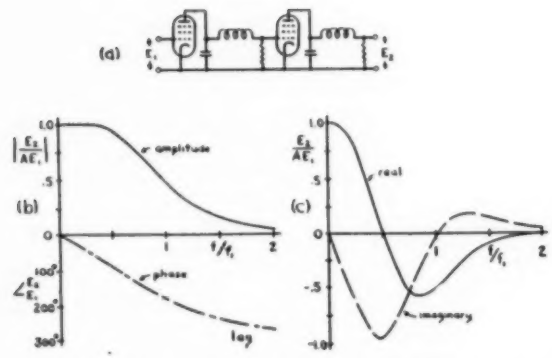


FIG. 5. (a) Low-pass amplifier circuit for use in examples. E_1 and E_2 are r.m.s. voltages of a single frequency. A is the zero-frequency voltage amplification. f_1 is a reference frequency. (b) Steady-state transmission characteristics of (a), amplitude and phase. (c) Steady-state transmission characteristics of (a), real and imaginary parts.

curves having different values of n. These points, where $m=n$, are uniquely determined by a single curve and, therefore, may be easily located.

IV. EXAMPLE FOR TYPICAL SOLUTIONS

As an example to illustrate the application of the graphical methods, the circuit of Fig. 5(a) will be used. This circuit consists of two identical sections of a low-pass network, coupled by pentode vacuum tubes. The r.m.s. voltage of a given frequency applied at the input of the system is E_1, and the resulting r.m.s. voltage at the output is E_2. For a particular choice of the component element values, the steady-state transmission characteristics of the system are shown in Fig. 5(b) and (c). Figure 5(b) gives these characteristics in terms of amplitude and phase, while Fig. 5(c) represents the same characteristics expressed as a real and an imaginary part. These parts are the products of the amplitude by the cosine and sine, respectively, of the phase angle. At a given frequency, the real part of the transmission characteristics is the ratio between the r.m.s. value of that component of the output voltage, which is in phase with the input voltage, and the r.m.s. value of the input voltage. Similarly, the imaginary part of the transmission characteristics is the corresponding ratio involving that component of the output voltage which leads the input voltage by ninety degrees. In the figure only positive frequencies are shown, although the shapes of the curves for negative frequencies must be understood. The amplitude curve and the real curve have even symmetry; the angle curve and the imaginary curve have odd symmetry. The scales of these curves are given as dimensionless ratios. The factor, A, is the voltage amplification of the system at zero frequency, and the reference frequency, f_1, is the frequency at which the inductor and capacitor of the circuits are in series resonance.

A non-sinusoidal input voltage, which has an instantaneous value, $e_1(t)$, is now applied to the system. In order not to complicate the example too much, this voltage is assumed to be an impulse function, having an infinite amplitude, zero duration, but a finite area equal to E_0t_0. The transform, $g_1(f)$, of such an impulse is easily shown to be a constant, E_0t_0, at all frequencies. This fact might be obtained from the curves of Fig. 2 and Fig. 3 for the rectangular wave shape, by allowing the duration, t_0, to approach zero as the amplitude, E_0, becomes infinite, like E_0/t_0. The product of this constant value multiplied by the steady-state transmission characteristics of the system therefore has the shape of these characteristics, with the ordinates merely changed by the constant factor, E_0t_0. This product is the frequency function, $g_2(f)$, at the output of the system. Its real part appears in Figs. 6(a), 7(a), and 8(a).

V. DIRECT CURVE-FITTING METHOD

Perhaps the most direct graphical method of obtaining the output signal, $e_2(t)$, is to make use of the curves of Figs. 2 and 3, combining them graphically so as to fit the frequency function in

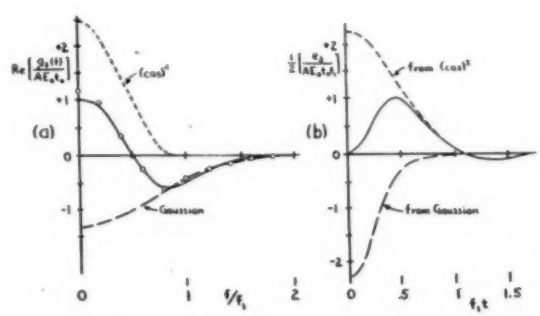


FIG. 6. Solution by direct curve fitting. (a) Real part of frequency function at output of system, as sum of simple curves. Circles indicate sum of dotted curves. (b) Output signal from system, as sum of inverse transforms of simple curves.

question. This procedure is illustrated in Fig. 6. The solid curve in Fig. 6(a) is the real part of the frequency function at the output of the system of Fig. 5(a), when an impulse of area E_0t_0 is applied at the input. Again, dimensionless ratios are used on the coordinate axes, and only the positive half of each curve is shown. This curve may be approximated by the sum of the cosine-squared curve and the (negative) Gaussian curve of Fig. 2, with their scales suitably adjusted. Although the curves of Fig. 2 are given as functions of time, it has already been pointed out that their arguments may be changed to frequency if desired. It is convenient to make the change for the present purpose. This curve-fitting is a trial and error process, but is not too difficult. The sum of the two curves is indicated by the circles in Fig. 6(a) and is seen to fit the given curve fairly well.

In Fig. 6(b) the transforms of the two standard frequency functions are plotted, having been obtained with the help of Fig. 3, and with careful observance of the coordinates. Again, only the positive half of each curve is shown. The solid curve of Fig. 6(b) is the sum of the two standard transforms. It has the proper shape for the output signal from the system, but the ordinates have just half the correct value. If the ordinates are doubled, this curve represents the desired output signal.

Where the shape of the frequency function is not too complicated, as is true here, it may be fitted reasonably well by a simple combination of two or three of the standard curves. After some practice this may be done with a minimum of trial and error. Obviously, the more closely the sum of the standard curves approximates the unknown curve, the more accurate will be the final results.

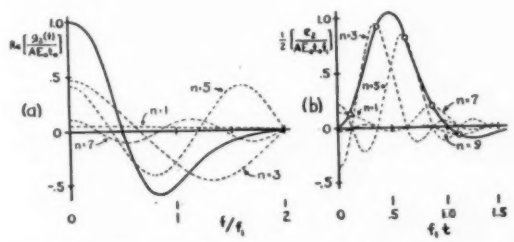


FIG. 7. Solution by cosine analysis of frequency function. (a) Real part of frequency function at output of system, as sum of cosines. (b) Output signal from system, as sum of transforms of cosines.

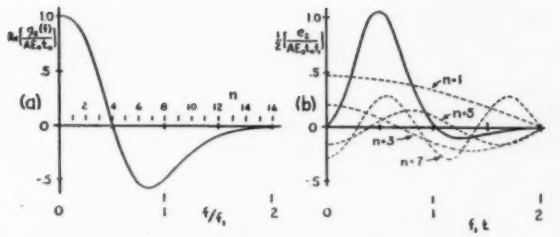


FIG. 8. Solution by cosine synthesis of time function. (a) Real part of frequency function at output of system, with selected points indicated. (b) Output signal from system, as sum of cosines.

VI. COSINE ANALYSIS OF THE FREQUENCY FUNCTION

In order to eliminate the trial and error of the type of curve-fitting just described, the cosine functions of Fig. 4 and their transforms may be used. It is necessary to interchange time and frequency in these curves, but since they are real even functions, the interchange may be made with no other modifications. The frequency function at the output of the system may be analyzed as a sum of cosine functions by any of a number of standard graphical methods. For example, the Fischer-Hinnen method[7] is well suited for this purpose. The transforms of the cosine functions may then be found from Fig. 4 and summed.

This process is illustrated in Fig. 7. The solid curve in Fig. 7(a) represents the real part of the frequency function at the output of the system, just as before, and is plotted in the same way. This curve is analyzed as the sum of cosines having odd numbers of quarter cycles in the frequency interval, $f_a = 2f_1$. This choice of this interval must be such that the given curve has essentially zero value for frequencies outside of the interval. The cosine functions thus obtained are plotted in Fig. 7(a) with the parameter, n, indicating the number of quarter cycles executed in the interval. Only values of n up to seven are plotted, although higher values are actually required to represent the curve well.

The transforms of each of these cosine functions may be found with the help of Fig. 4. The transforms are plotted in Fig. 7(b), showing only values of n up to nine, and are summed. Once

[7] F. A. Laws, *Electrical Measurements* (McGraw-Hill Book Company, Inc., New York, 1938), second edition, pp. 687–695.

again, if the ordinates of this sum are doubled, the output signal from the system is obtained. The larger the number of cosines used in the original analysis, the more accurately they will approximate the frequency function, and the more accurate will be the final result.

This analysis of Fig. 7 was carried out with n chosen odd, and the curves of Fig. 4(c) and (d) were used. Exactly the same results will be obtained if n is chosen even, although the analysis is different and the curves of Fig. 4(a) and (b) would be employed. There is usually little basis for choosing between these two analyses since either will produce the same result with about the same amount of effort.

The shape of the transforms in Fig. 4(b) and (d) is relatively complicated, and their plotting and summing as in Fig. 7(b) is tedious. However, if it is sufficient to obtain only a few discrete points on the curve of the output signal, the properties of these functions are such that their plotting may be eliminated. It was pointed out earlier, in connection with Eq. (4), that there are certain times at which only the curve for a single value of n contributes to the final sum. These points are shown by circles in Fig. 7(b). The times at which these points occur are $n/(4f_a)$, and the values at these times are $a_n f_a$, where a_n is the amplitude of the nth cosine function of the analysis, and n is not zero. If n is zero, the value at $t=0$ is $2a_0 f_a$. These points are located by data from the cosine analysis, and plotting of the transforms is not required. Evidently, if it is desired to obtain a large number of points closely spaced along the curve of the output function, the frequency interval, f_a, must be chosen relatively large and the cosine analysis carried to large values of n.

VII. COSINE SYNTHESIS OF TIME FUNCTION

A further simplification in the analysis results if the foregoing procedure is applied in reverse order. The unknown frequency function is analyzed as the sum of functions having the shape of those of Fig. 4(b) and (d), and their transforms then found as cosine functions. In this case, the frequency and time scales are left as in Fig. 4. This type of analysis is simpler than the analysis into cosines, since it involves only a direct reading of the value of the curve at certain points. It is not necessary to plot the curves into which the frequency function is analyzed.

This process is illustrated in Fig. 8, where once more Fig. 8(a) represents the real part of the frequency function at the output of the system. The abscissa of this curve is divided into frequency intervals, f_b, chosen in this case as $f_1/8$. These intervals are numbered consecutively as shown, with n equal to zero at zero frequency. The value of the curve at points corresponding to the various values of n may then be read directly. The transform of the curve is the sum of the ordinates of a number of cosine curves which execute n quarter cycles in the time interval, $t_b=1/(4f_b)=2/f_1$, where n is chosen either odd or even. The amplitude of the nth cosine is p_n/t_b where p_n is the value of the frequency function at nf_b, and n is not zero. If n is zero, the amplitude of the zeroth cosine (i.e., constant term) is $p_0/(2t_b)$. Cosine time functions, with n odd and values only up to seven, are plotted in Fig. 8(b) and summed. Their sum, with its ordinates doubled, is the output signal from the system. The same result would have been obtained had n been chosen even instead of odd, and once more there is little basis for choice in the matter. If the final time function is to be relatively accurate, the frequency interval, f_b, must be chosen small enough so that information is obtained from each important part of the frequency function.

VIII. INPUT SIGNAL IN THE FORM OF A STEP FUNCTION

In the preceding examples, the input signal to the system has been chosen as an impulse function, since this function has the simplest possible transform. For many applications, in the analysis of the behavior of servomechanisms[8] for instance, the step function is often used for test purposes. The step function, $e_s(t)$, is defined as

$$\begin{cases} e_s(t)=0 & \text{for } t<0 \\ e_s(t)=E_0 & \text{for } t>0. \end{cases}$$

The step function may also be written $e_s(t)$

[8] H. Lauer, R. Lesnick, L. E. Matson, *Servomechanism Fundamentals* (McGraw-Hill Book Company, Inc., New York, 1947), Chap. IV.

$= E_0/2 + e_s''(t)$ where $e_s''(t)$ is defined as

$$\begin{cases} e_s''(t) = -E_0/2 & \text{for} \quad t<0 \\ e_s''(t) = +E_0/2 & \text{for} \quad t>0 \end{cases}$$

and is a real odd function of time. The usual step function, $e_s(t)$, is thus replaced by a constant, $E_0/2$, and the symmetrical step, $e_s''(t)$. The transform of $e_s''(t)$ may be found from Eq. (1) as $g_s''(f) = -jE_0/\omega$, and is an imaginary odd function of frequency. This result is obtained by multiplying the time function by factors of the form $\exp(\pm ct)$, where c is a real constant, the sign of which is chosen so as to make the product vanish at plus and minus infinity. After completing the integration of Eq. (1), c is allowed to become zero, whereupon the transform given above is found.

The response of a system to the symmetrical step function, $e_s''(t)$, may be determined by any of the methods already described in connection with the impulse function. The response to the constant, $E_0/2$, is simply the product of A (the amplification of the system at zero frequency) and $E_0/2$. Finally, the sum of the responses to the symmetrical step, $e_s''(t)$, and to the constant, $E_0/2$, is the response of the system to the step function, $e_s(t)$. If only the real part of the frequency function at the output is used in the solution, the ordinates of the result must be doubled. For a passive system, at rest at $t=0$, the response must also be zero at $t=0$. After a long time the response must be AE_0.

IX. FINAL REMARKS

In each of the three graphical methods described here only the real part of the frequency function at the output of the system was utilized. It would be possible, of course, to make the analysis in terms of the imaginary part alone. There are occasions when both the real and the imaginary parts of the function must be considered. This is true, for example, if a real time function having no symmetry is to be transformed into its corresponding frequency function. In such a case, the frequency function will be complex, and both real and imaginary parts must be found. The necessary transformations may follow any of the procedures already described, since the same methods allow transformation from frequency to time, or *vice versa*.

The imaginary part may be handled most easily by a procedure analogous to the third method described, in which the real part was treated in terms of a sum of cosine functions. Inasmuch as the imaginary part must have odd symmetry, it must be treated in terms of a sum of sine functions. It is not difficult to set up families of sine functions and their transforms analogous to the cosine functions of Fig. 4. A negative sine curve has as a transform a curve similar in shape to those of Fig. 4(b) and (d), with the main hump positive. The curves necessarily have odd symmetry, and one function is imaginary while the other is real. The most important fact so far as an analysis in sine functions is concerned, is that the relations expressed in Eq. (4) for the transforms of cosine functions are also valid for the transforms of sine functions. Therefore, the analysis in terms of sines proceeds in the same way as has been discussed here in terms of cosines. The only differences are the opposite type of symmetry and the negative algebraic sign which must appear in such a ratio as $-p_n/t_b$, giving the amplitude of the nth sine for synthesis by the third method.

Only the very simple cases of an impulse function and a step function at the input of the system were considered here. If the input signal is more complicated so that its transform is not readily calculable, a graphical analysis of the type already described may be used to find this transform. If the input signal is real, but has no symmetry, steps (d) through (i) of Fig. 1 must be carried out, in reverse order, of course, to yield the frequency function, $g_1(f)$, corresponding to the input signal. With this function known, the remainder of the analysis proceeds as already described.

Of the three graphical methods discussed here, the third one in which the output signal is synthesized as a sum of cosine (or sine) functions is probably the most useful. However, if the curves are relatively simple, the first method, involving an experimental fitting of the curves, may lead more quickly to a result. As in all graphical methods, the accuracy of the results obtained from any of the procedures depends largely upon the skill with which the plotting is carried out. If only the approximate shape of the

output signal is required, the graphical curve-fitting may be simplified so that it may be done in a short time. If accurate results are needed then more careful analysis is required.

Finally, it should be pointed out that the third method is in many problems essentially the same as that used by Bedford and Fredendall.[4] However, the bases for the two methods are somewhat different, and only in certain cases do the two become identical. In the present method, the frequency function at the output of the system is considered in terms of its real or imaginary parts, leading to the synthesis of the output signal as a sum of cosines or sines, all of which have zero relative phase displacement. In the method used by Bedford and Fredendall, the output signal is synthesized as a sum of cosines which have phase displacements determined by the problem. For a given number of components, a better representation may be obtained by the latter method, since an additional variable is allowed. On the other hand, the manipulation required to obtain the result may be simpler in the method described here.

X. ACKNOWLEDGMENT

The original work reported here was done under the able guidance of Professor E. L. Chaffee and his helpful comments and suggestions are gratefully acknowledged.

Two Calculating Machines for X-Ray Crystal Structure Analysis

A. D. Booth

Institute for Advanced Study, Princeton, New Jersey, and Birkbeck College, England

(Received July 2, 1946)

The calculation of structure factors in x-ray structure analysis is one of the most laborious operations. Two mechanisms are described which have proved of great service in several analyses. The simpler of the devices is of general application and can be constructed with comparatively limited workshop facilities.

INTRODUCTION

APART from large and expensive machines for performing automatically the whole range of crystallographic calculation, there remains a need for simple *ad hoc* devices to deal with particular aspects of the problem. If the current schemes for centralizing the latter stages of Fourier refinement come to fruition, this demand for the "home made" and simple type of calculator is likely to be increased.

It is the purpose of this paper to describe two such arrangements which the author designed and found useful in practical structure analysis. The first is especially valuable in the case of tetragonal space groups and found extensive application during the determination of the structure of pentaerythritol tetranitrate.[1] The second is considerably simpler in principle and in design and has been in constant and satisfactory use in structure analysis[2,3] for the past four years.

THE TETRAGONAL CALCULATOR

In tetragonal space groups it is necessary to obtain the value of expressions of the type:

$$F(hkl)=\sum_r f_r\left[\cos 2\pi\frac{hx_r}{a}\cdot\cos 2\pi\frac{ky_r}{b}-\cos 2\pi\frac{kx_r}{a}\cdot\cos 2\pi\frac{hy_r}{b}\right]\cos 2\pi\frac{lz_r}{c} \quad (1)$$

for a range of values of (h, k, l). The information given being the values of $(hx_r/a,\ ky_r/b,\ lz_r/c)$ and f_r.

This calculation is performed by the mechanism shown in Fig. 1.

L_1N_1 and L_2N_2 are two bars of $\frac{3}{8}''$ diam. steel

[1] Booth and Llewellyn, Proc. Roy. Soc. (to be published).

[2] G. A. Jeffrey, Proc. Roy. Soc. **A183**, 388 (1945).

[3] G. A. Jeffrey, Proc. Roy. Soc. **A188**, 222 (1947).

rod which are capable of rotation in a horizontal plane about vertical axes through their centers, O_1 and O_2, respectively. Sliding along these bars are riders P_1 and P_2, so arranged as to be capable of fixation at any given radii OP_1 and OP_2 (positive or negative). Attached to the riders is a link mechanism VP_1RP_2; this has the bars P_1RP_2, VR rigidly fixed at right angles, and the links VP_1, VP_2 of equal length. The bar P_1RP_2 can slide freely through the rider points P_1P_2, whilst the links P_1V, P_2V can rotate about P_1, P_2, and V with the latter freely sliding on RV.

The effect of this mechanism is seen to be the maintenance of R at the mid point of P_1 and P_2.

If now the bars L_1N_1, L_2N_2 are rotated through angles θ_1 and θ_2 from the horizontal line O_1O_2, as shown, the displacement of R, parallel to O_1O_2, is simply:

$$\tfrac{1}{2}(O_2P_2\cos\theta_2 - O_1P_1\cos\theta_1). \qquad (2)$$

This component of the motion of R is separated by means of the framework $ABCD$ which can move, on rollers, in a horizontal line parallel to O_1O_2, R moving in a groove in AB.

The motion of CD is communicated, *via* a silk string S, to a disk E which rotates about a vertical axis through its center M. The top surface of E is covered with fine calico and has applied to it a planimeter assembly P. If the axis of the planimeter makes an angle θ_3 with the radius Mp through its point of contact p, the resultant rotation is

$$\Theta \cdot Mp \cdot \cos\theta_3, \qquad (3)$$

where Θ is the angle of rotation of E.

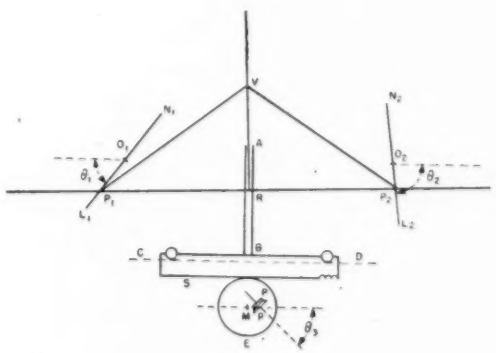


FIG. 1.

Thus, if scale factors are properly adjusted, the motion recorded on the planimeter is:

$$Mp(O_2P_2\cos\theta_2 - O_1P_1\cos\theta_1)\cos\theta_3. \qquad (4)$$

The planimeter assembly is provided with a clutch so that it can be removed from contact with E.

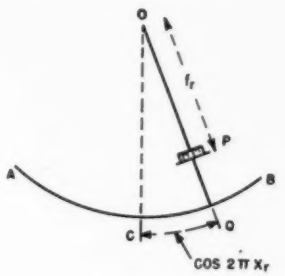


FIG. 2.

It follows, that by making:

$$Mp = f_r$$

$$O_2P_2 = \cos 2\pi\frac{hx_r}{a},$$

$$\theta_2 = \cos 2\pi\frac{ky_r}{b},$$

$$O_1P_1 = \cos 2\pi\frac{kx_r}{a},$$

$$\theta_1 = \cos 2\pi\frac{hy_r}{b},$$

$$\theta_3 = 2\pi\frac{lz_r}{c},$$

and resetting between operations, the summation (1) can be computed. The bars L_1N_1, L_2N_2 bear upon their upper surfaces scales directly calibrated in $\cos 2\pi(hx/a)$.

The machine proved most effective when two operators were available, one for each turntable, L_1N_1, L_2N_2, and reduced the time of computing the complicated expression (1) by a factor of approximately 4.

Unfortunately no photographs of the machine were taken and it has now, with the ending of the war, been dismantled.

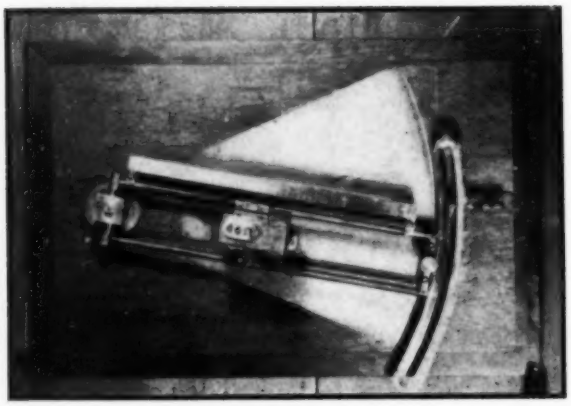


FIG. 3.

THE COSINE ADDING MACHINE

This is considerably simpler in principle than the machine just described and has the advantage of not requiring any delicate mechanism of the planimeter type.

Although not always the most mathematically elegant method, it is easily shown that any structure factor formula can be reduced to a set of summations of the type:

$$\sum_r f_r \left\{ \begin{matrix} \sin \\ \cos \end{matrix} \right\} 2\pi \left(\frac{hx_r}{a} \pm \frac{ky_r}{b} \pm \frac{lz_r}{c} \right). \qquad (5)$$

A table of the functions $hx_r/a \pm ky_r/b + lz_r/c$ is readily computed using a "Facit" or other multiplying machine so that it is desirable to have a simple and rapid means of effecting the summation:

$$\sum_r f_r \left\{ \begin{matrix} \sin \\ \cos \end{matrix} \right\} 2\pi X_r. \qquad (6)$$

This is found in the mechanism shown in Fig. 2. A planimeter wheel P is mounted on a bar OQ pivoted at O: if the bar is turned through an angle $CQ = K \cos 2\pi X_r$, where K is constant, the motion recorded by P is simply:

$$K' \cdot OP \cdot \cos 2\pi X_r.$$

The constants are so arranged that $OP \propto f_r$ so that:

$$f_r \cos 2\pi X_r$$

is recorded.

By removing P from contact with the plane $OABC$ during the return of Q to its zero, C, and resetting OP, the complete summation (6) can be rapidly effected.

The machine is particularly advantageous in organic structure analysis where f_r is constant for a large number of atoms.

Practical details are shown in Fig. 3. The planimeter wheel has a vulcanized rubber tire, the outside diameter being 1.625". The counter mechanism was extracted from an old electricity meter with a step up ratio between wheel and counter of 1:7. The angle scale of $\cos 2\pi X_r$ is in the form of a circular segment and is attached to the planimeter on rail assemblies as shown. A simple spring push type clutch is provided so that the planimeter is normally out of contact with the friction plate, the latter being of sand blasted brass. An audible "click" is given on the return of the setting to zero.

This machine proves not much inferior in speed to a normal key driven adding machine.

DISCUSSION

The second, and simpler of these machines, has been and is of constant utility and is sufficiently versatile to be worth making a permanent fixture in a crystallographic laboratory. The first is of more ephemeral application, but is worthy of description as containing a number of mechanical elements which can be incorporated in other special purpose machines, and which may consequently be of service to other workers in the field.

The author wishes to record his thanks to Mr. H. W. Giles for his care and numerous helpful suggestions while constructing the second machine, and to the Board of the British Rubber Producers' Research Association under whose auspices the machine was constructed. For the first machine valuable mechanical help was given by Mr. W. H. Small, B.Sc.

The Field of a Microwave Dipole Antenna in the Vicinity of the Horizon*

C. L. PEKERIS**
Columbia University Mathematical Physics Group, New York, New York
(Received April 8, 1947)

In this paper a method is developed for determining the electromagnetic field produced by a microwave antenna at points on the horizon, and on either side of it, where neither the ray theory nor the normal mode theory can be used conveniently. The theory is developed for a condition of standard atmospheric refraction, by use of a space in which the earth is flattened and the rays are curved. This allows us to make a simple derivation of the ray theory, valid in the optical region, and of the normal mode theory, suited for the shadow zone. For the intermediate region centered around the horizon we use the original integral for the potential to obtain expressions for the field under the restriction of maximum absorption, which for typical ground conditions applies to wave-lengths less than about a meter. Three cases are treated in which the transmitter, or receiver, are either situated on the ground or are elevated several natural units of height. For an elevated transmitter and receiver the Hertzian potential Ψ due to a point source at the origin is in the vicinity of the horizon given by

$$|\Psi| = [1/2(r\hat{r})^{\frac{1}{2}}]\left|(2/\pi^{\frac{1}{2}})e^{i\pi/4}\int_0^\infty \exp[-i(2\tau^{\frac{1}{2}}t+t^2)]dt - [(2/3)^{\frac{1}{2}}\Omega/\pi^{\frac{1}{2}}]e^{-i\pi/12}F(p)\right|, \quad \text{(A)}$$

with

$$p=(3/2)^{\frac{1}{3}}(x-\hat{x}),\ \Omega^2=(z_1^{-\frac{1}{2}}+z_2^{-\frac{1}{2}}),$$
$$\tau=(x-\hat{x})^2/\Omega^2,\ \hat{x}=z_1^{\frac{1}{2}}+z_2^{\frac{1}{2}},$$

where x denotes the horizontal distance r expressed in natural units, z_1 and z_2 the heights of transmitter and receiver in natural units, $\hat{r}$ the distance of receiver from transmitter when the former is on the horizon. $F(p)$ (see Eq. (68) below) has been evaluated, and is given in Table IV, while the integral in (A) can be expressed in terms of the tabulated Fresnel integrals. In the limit of very short wave-lengths the field *on* the horizon approaches the value $1/(2\hat{r})$ which would result from the diffraction of the *direct ray only* by a straight edge placed at the point of tangency of the horizon with the earth. A comparison of the field obtained from (A) with exact values computed by van d. Pol and Bremmer, using the ray theory and the normal mode theory, is shown in Figs. 6 and 7.

When the transmitter is at zero elevation and the receiver is elevated several units of height, the potential in the vicinity of the horizon is given by

$$|\Psi_v| = \frac{(3/2\pi a_e)^{\frac{1}{2}}\lambda^{\frac{1}{2}}}{\pi(r\hat{r})^{\frac{1}{2}}}\left|\frac{G(p)\epsilon_1}{(\epsilon_1-1)^{\frac{1}{2}}}\right|, \quad \text{(B)}$$

$$|\Psi_h| = \frac{(3/2\pi a_e)^{\frac{1}{2}}\lambda^{\frac{1}{2}}}{\pi(r\hat{r})^{\frac{1}{2}}}\left|\frac{G(p)}{(\epsilon_1-1)^{\frac{1}{2}}}\right| \quad \text{(C)}$$

for vertical polarization and horizontal polarization, respectively. Here ϵ_1 denotes the complex dielectric constant, a_e the effective radius of the earth, and λ the wave-length. $G(p)$ is given in Eq. (78) and is shown in Fig. 4. A comparison of (B) with exact values obtained by van d. Pol and Bremmer is shown in Fig. 8.

When both the transmitter and receiver are at zero elevation, it is found that the potential can be expressed as the *sum* of the surface wave appropriate for a *flat* ground and an integral depending on the radius of the earth. At great distances, the two terms tend to cancel out. Under conditions of maximum absorption this leads to

$$|\Psi_v| = \frac{2}{r^2k_0}\left|\frac{\epsilon_1^2 g(p')}{(\epsilon_1-1)}\right|;\quad |\Psi_h| = \frac{2}{r^2k_0}\left|\frac{g(p')}{(\epsilon_1-1)}\right|, \quad \text{(D)}$$

$$p'=(3/2)^{\frac{1}{3}}x,\quad g(p')=1-(2e^{i\pi/4}/3\pi^{\frac{1}{2}})(p')^{\frac{1}{2}}H(p').$$

$H(p')$ is given in (84) and $g(p')$ is shown in Fig. 5, where it is compared with results obtained previously by van d. Pol and Bremmer using the normal mode theory.

For points *on* the horizon Eq. (A) reduces to

$$|\Psi| = (1/2\hat{r})|1-(0.684-0.183i)\Omega|, \quad \text{(E)}$$

while in (B) and (C) we put $|G(0)|=2.13$.

1. INTRODUCTION

IN the treatment of the problem of diffraction of radio waves around the earth[1] there are available at present the ray method, which is valid only in the lit region, and the method of normal modes, which is suited for the shadow region beyond the horizon. On the horizon and beyond the ray method breaks down completely, while the method of normal modes requires an increasing number of modes to be used as one recedes from the horizon into the lit

* This paper is based on work done for the Naval Research Laboratory under contract with the Navy's Office of Research and Inventions. Publication assisted by the Ernest Kempton Adams Fund for Physical Research of Columbia University.

** On leave of absence from Massachusetts Institute of Technology.

[1] G. N. Watson, Proc. Roy. Soc. **A95**, 83 (1918). B. van d. Pol and H. Bremmer, Phil. Mag. **24**, 141, 825 (1937); **25**, 817 (1938); **27**, 261 (1939). B. Wwedensky, Tech. Phys. U.S.S.R. **2**, 624 (1935); **3**, 915 (1936); **4**, 579 (1937). T. L. Eckersley and G. Millington, Phil. Trans. Roy. Soc. **237**, 273 (1938). P. S. Epstein, Proc. Nat. Acad. Sci. **21**, 62 (1935). G. Millington, Phil. Mag. **27**, 517 (1939). C. R. Burrows and M. C. Gray, Proc. I.R.E. **29**, 16, 1941. K. A. Norton, F. C. C. Report 39920 (1940); F. C. C. Report 47475 (1941).

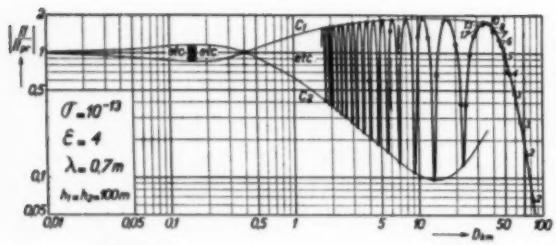


FIG. 1. Black circle points were calculated by the ray theory. The numbers placed near the open circles represent the number of modes that had to be included in order to obtain an accuracy of one percent. The horizon is at the dashed line (van d. Pol and Bremmer, Phil. Mag. **27**, 270 (1939)).

region. This is illustrated in Fig. 1 which is taken from the work of van d. Pol and Bremmer.[2] The purpose of this investigation is to develop a theory of diffraction of microwaves which is intrinsically most accurate in the vicinity of the horizon, and which becomes less accurate as one recedes from the horizon either far into the lit zone or far into the shadow zone, thus bridging the gap between the ray theory and the normal mode theory. The horizon is the locus of coalescence of the direct ray with the reflected ray, and for this reason our work displays some points of similarity with Airy's theory of diffraction near a caustic. Expression (A) for the Hertzian potential given above reduces to an interference pattern between the direct and reflected rays for points inside the lit zone (Figs. 6 and 7), while inside the shadow zone it is a monotone decreasing function caused by cancellation of the two terms.

The propagation of microwaves around the earth presents a peculiar diffraction problem, in that the distance of the source from the sphere, as well as the wave-length, is small in comparison with the radius of the sphere. Furthermore, we are interested in the diffracted field at points which are within a range of only a fraction of the radius from the sphere. The question has often been raised to what extent an object of finite curvature diffracts like a straight edge placed at the point of tangency of the horizon line. In this connection our analysis shows that, within ranges of the order of a fraction of the radius from the sphere, the field *on the horizon* approaches the value for a straight edge in the limit of short wave-lengths, provided we neglect the contribution from the reflected ray. The latter is justified because of the divergence of the reflected ray near the horizon.

Our treatment differs from Watson's, and its later extensions by van d. Pol and Bremmer,[1] in that the problem of diffraction around the spherical earth is transformed into one of diffraction over a flat earth through an atmosphere with a suitably variable index of refraction. This device was first introduced by Schelleng, Burrows and Ferrell,[3] and it is known to be valid[4] out to ranges of the order of the radius of the earth and for wave-lengths greater than a centimeter up to heights of the order of a thousand feet. By this method one obviates the synthesis of an original solution expressed as a series of spherical harmonics, and, furthermore, arrives directly at an integral for the potential containing Bessel functions of the order $\frac{1}{3}$, which are intrinsic to the problem.

The effect of the curvature of the earth is introduced by replacing the actual index of refraction μ by a modified index N defined by

$$N(r)=r\mu(r)/a\mu(a), \tag{1}$$

a denoting the radius of the earth and r the distance from the center of the earth. In a free atmosphere $N(r)$ increases with height h above the surface at the rate of h/a, while under conditions of standard atmospheric refraction this rate is changed to h/a_e, with[3] $a_e \simeq 4a/3$. Our problem reduces therefore to finding a solution for the Hertzian potential Ψ satisfying the wave equation:

$$\nabla^2\Psi+k_0^2\mu^2\Psi=0, \quad k_0=2\pi/\lambda, \tag{2}$$

$$\mu^2=(1+qh), \quad q=2/a_e, \quad \text{in air}, \tag{3}$$

$$\mu^2=(1+qh)\epsilon_1, \quad \epsilon_1=\epsilon-i\sigma/\omega, \quad \text{in the ground}, \tag{4}$$

where ϵ denotes the dielectric constant and σ the conductivity of the ground, and a time factor $e^{i\omega t}$ has been assumed. It will be noted that it is necessary to distort the space inside the earth as well. The fact that the expression for μ^2 in (4) becomes formally negative inside the earth (where h is negative) at a depth of about half a_e

[2] B. van d. Pol and H. Bremmer, Phil. Mag. **27**, 270, (1939).

[3] J. C. Schelleng, C. R. Burrows, and E. B. Ferrell, Proc. I.R.E. **21**, 427 (1933); J. E. Freehafer, Radiation Laboratory Report 447 (1943).

[4] C. L. Pekeris, Phys. Rev. **70**, 518 (1946).

need not concern us, since the earth-flattening approximation is to be applied only under conditions where the energy inside the earth is confined to a thin surface layer.

It should be noted that the problem treated here finds an application in underwater acoustics,[5] where the index of refraction μ for sound propagation sometimes varies with depth h according to Eq. (3). In that problem of a real flat boundary and a variable μ, the parameter q in (3), which is a measure of the curvature of the rays, need not be small. On the other hand, in the application to propagation of microwaves around the earth, the earth-flattening approximation is valid only insofar as powers of qh higher than the first can be neglected, so that in the sequel, the expression for the field should be relied on only as far as the first power in qh.

2. FORMAL SOLUTION

The solution of (2) must reduce to the form e^{-ikR}/R at small distances R from the source, it must represent an outgoing wave at great distances from the source, and the tangential components of the electric and magnetic fields must be continuous at the surface of the earth. The formal solution of (2), satisfying the above mentioned boundary conditions, can be obtained by a method given by H. Lamb.[5,6] The solution is built up from elementary solutions of the form

$$\psi = e^{i\omega t} J_0(kr) F(h) G(k), \tag{5}$$

r now denoting horizontal distance, and k a variable over which one eventually integrates in the complex k-plane. $F(h)$ satisfies the equations

$$(d^2F/dh^2) + [k_0^2(1+qh) - k^2]F = 0, \quad \text{in the air } (h>0), \tag{6}$$

$$(d^2F/dh^2) + [k_1^2(1+qh) - k^2]F = 0; \quad k_1^2 = k_0^2\epsilon_1, \quad \text{in the ground } (h<0). \tag{7}$$

The fundamental solutions of (6) are

$$M(u) = u^{\frac{1}{2}} H_{1/3}^{(1)}(2u^{\frac{3}{2}}/3), \quad N(u) = u^{\frac{1}{2}} H_{1/3}^{(2)}(2u^{\frac{3}{2}}/3), \quad u = (k_0/q)^{\frac{2}{3}}[1+qh-(k^2/k_0^2)], \tag{8}$$

while the solution of (7), which does not become infinite at great depth, is

$$P(v) = v^{\frac{1}{2}}[J_{1/3}(2v^{\frac{3}{2}}/3) + J_{-1/3}(2v^{\frac{3}{2}}/3)] = [(3v)^{\frac{1}{2}}/2][H_{1/3}^{(1)}(2v^{\frac{3}{2}}/3)e^{i\pi/6} + e^{-i\pi/6}H_{1/3}^{(2)}(2v^{\frac{3}{2}}/3)],$$
$$v = (k_1/q)^{\frac{2}{3}}[1+qh-(k^2/k_1^2)]. \tag{9}$$

Here the phase of u is zero when positive, and $-i\pi$ when negative, and similarly for v. Let the transmitter's height above the ground be denoted by h_1, then we have

$$\begin{aligned} F_1(h) &= C(k)N(u), \quad h>h_1, \\ F_2(h) &= A(k)M(u) + B(k)N(u), \quad 0<h<h_1, \\ F_3(h) &= E(k)P(v), \quad h<0, \end{aligned} \tag{10}$$

where the functions of k, A, B, C, and E are determined by the boundary conditions. These are, for a point source at h_1,

$$F_1 = F_2, \quad (\partial F_2/\partial h) - (\partial F_1/\partial h) = 2k, \quad \text{at } h = h_1, \tag{11}$$

$$\partial F_2/\partial h = \partial F_3/\partial h, \quad k_0^2 F_2 = k_1^2 F_3 \quad \text{at } h=0, \text{ for vertical polarization,} \tag{12}$$

$$\partial F_2/\partial h = \partial F_3/\partial h, \quad F_2 = F_3 \quad \text{at } h=0, \text{ for horizontal polarization.} \tag{13}$$

Using the relations

$$dM/du = uH_{-2/3}^{(1)}(2u^{\frac{3}{2}}/3), \quad dN/du = uH_{-2/3}^{(2)}(2u^{\frac{3}{2}}/3), \tag{14}$$

$$H_\nu^{(1)}(z)\cdot H_{\nu+1}^{(2)}(z) - H_{\nu+1}^{(1)}(z)\cdot H_\nu^{(2)}(z) = 4i/\pi z, \tag{15}$$

[5] C. L. Pekeris, J. Acous. Soc. Am. **18**, 295 (1946).
[6] H. Lamb, Phil. Trans. Roy. Soc. **A203**, 1 (1904); W. H. Furry, Radiation Laboratory Report 680 (1945). See also Arnold Sommerfeld's article in Frank-Mises, *Differential-gleichungen der Physik* (Rosenberg, New York, 1943) Vol. II, p. 918.

in solving for A, B, C, and E from (11) and (12), or from (11) and (13), we obtain

$$\Psi_1 = -[i\pi e^{i\omega t}/3(qk_0^2)^{\frac{1}{2}}]\int_0^\infty J_0(kr)kdk(u_1u)^{\frac{1}{2}}H_{1/3}^{(2)}(2u^{\frac{3}{2}}/3)[H_{1/3}^{(1)}(2u_1^{\frac{3}{2}}/3)+ \\ +(B/A)H_{1/3}^{(2)}(2u_1^{\frac{3}{2}}/3)], \quad h>h_1, \quad (16)$$

$$\Psi_2 = -[i\pi e^{i\omega t}/3(qk_0^2)^{\frac{1}{2}}]\int_0^\infty J_0(kr)kdk(u_1u)^{\frac{1}{2}}H_{1/3}^{(2)}(2u_1^{\frac{3}{2}}/3)[H_{1/3}^{(1)}(2u^{\frac{3}{2}}/3) \\ +(B/A)H_{1/3}^{(2)}(2u^{\frac{3}{2}}/3)], \quad 0<h<h_1, \quad (17)$$

$$\Psi_3 = -[i\pi\gamma e^{i\omega t}/3(qk_0^2)^{\frac{1}{2}}]\int_0^\infty J_0(kr)kdk[P(v)/P(v_0)]H_{1/3}^{(2)}(2u^{\frac{3}{2}}/3)[H_{1/3}^{(1)}(2u_0^{\frac{3}{2}}/3)+ \\ +(B/A)H_{1/3}^{(2)}(2u_0^{\frac{3}{2}}/3)], \quad h<0, \quad (18)$$

$$u_0=(k_0/q)^{\frac{1}{2}}[1-(k^2/k_0^2)], \quad u_1=(k_0/q)^{\frac{1}{2}}[1+qh_1-(k^2/k_0^2)], \quad v_0=(k_1/q)^{\frac{1}{2}}[1-(k^2/k_1^2)], \quad (19)$$

$$B/A=-[\beta u_0^{\frac{1}{2}}H_{1/3}^{(1)}(2u_0^{\frac{3}{2}}/3)-u_0H_{-2/3}^{(1)}(2u_0^{\frac{3}{2}}/3)]/[\beta u_0^{\frac{1}{2}}H_{1/3}^{(2)}(2u_0^{\frac{3}{2}}/3)-u_0H_{-2/3}^{(2)}(2u_0^{\frac{3}{2}}/3)], \quad (20)$$

$$\beta=(1/\epsilon_1^{\frac{1}{2}})[\dot{P}(v_0)/P(v_0)]\simeq(iv_0^{\frac{1}{2}}/\epsilon_1^{\frac{1}{2}}); \quad \gamma=1/\epsilon_1, \text{ for vertical polarization}, \quad (21)$$

$$\beta=\epsilon_1^{\frac{1}{2}}[\dot{P}(v_0)/P(v_0)]\simeq iv_0^{\frac{1}{2}}\cdot\epsilon_1^{\frac{1}{2}}; \quad \gamma=1, \text{ for horizontal polarization}. \quad (22)$$

The only assumption made in deriving this solution, aside from the one involved in the earth-flattening approximation, was in writing $iv_0^{\frac{1}{2}}$ for $\dot{P}(v_0)/P(v_0)$. This approximation holds with increasing accuracy the more absorptive the ground is; even for dry soil the error introduced by this approximation is less than one part in 10^7.

We shall now show that in the limit of an infinite radius of the sphere ($q\to 0$) our solution reduces to Sommerfeld's solution[6] for the potential of a dipole over a *flat* earth. As $q\to 0$, u_0 defined in (19) becomes very large, and

$$u_1^{\frac{1}{2}}\to u_0^{\frac{1}{2}}+(3h_1/2)(k_0^2-k^2)^{\frac{1}{2}}+\cdots; \quad u\to u_0^{\frac{1}{2}}+(3h/2)(k_0^2-k^2)^{\frac{1}{2}}+\cdots. \quad (23)$$

We can therefore replace the Hankel functions in (16) and (17) by their asymptotic expansions:

$$H_{1/3}^{(1)}(z)\to(2/\pi z)^{\frac{1}{2}}\exp[i(z-5\pi/12)], \quad H_{1/3}^{(2)}(z)\to(2/\pi z)^{\frac{1}{2}}\exp[i(-z+5\pi/12)], \quad (24)$$

$$B/A\to\exp[-i5\pi/6+4iu_0^{\frac{3}{2}}/3]R(k), \quad k<k_0, \quad (25)$$

$$R(k)=\left[\frac{k_1^2(k_0^2-k^2)^{\frac{1}{2}}-k_0^2(k_1^2-k^2)^{\frac{1}{2}}}{k_1^2(k_0^2-k^2)^{\frac{1}{2}}+k_0^2(k_1^2-k^2)^{\frac{1}{2}}}\right], \text{ for vertical polarization},$$

$$=\left[\frac{(k_0^2-k^2)^{\frac{1}{2}}-(k_1^2-k^2)^{\frac{1}{2}}}{(k_0^2-k^2)^{\frac{1}{2}}+(k_1^2-k^2)^{\frac{1}{2}}}\right], \text{ for horizontal polarization}, \quad (26)$$

$$(u_1u)^{\frac{1}{2}}H_{1/3}^{(2)}(2u^{\frac{3}{2}}/3)[H_{1/3}^{(1)}(2u_1^{\frac{3}{2}}/3)+(B/A)H_{1/3}^{(2)}(2u_1^{\frac{3}{2}}/3)] \\ \to\frac{3(qk_0^2)^{\frac{1}{2}}}{\pi(k_0^2-k^2)^{\frac{1}{2}}}\{\exp[-i(h-h_1)(k_0^2-k^2)^{\frac{1}{2}}]+R(k)\exp[-i(h_1+h)(k_0^2-k^2)^{\frac{1}{2}}]\}, \quad (27)$$

$$\Psi_1\to -i\int_0^\infty\frac{J_0(kr)kdk}{(k_0^2-k^2)^{\frac{1}{2}}}\{\exp[-i(h-h_1)(k_0^2-k^2)^{\frac{1}{2}}]+R(k)\exp[-i(h_1+h)(k_0^2-k^2)^{\frac{1}{2}}]\}, \quad h>h_1. \quad (28)$$

This agrees with Sommerfeld's solution for a flat earth in which $h_1\to 0$, since $(k_0^2-k^2)^{\frac{1}{2}}\to -i(k^2-k_0^2)^{\frac{1}{2}}$ for $k>k_0$.

[7] Reference 5, p. 296.

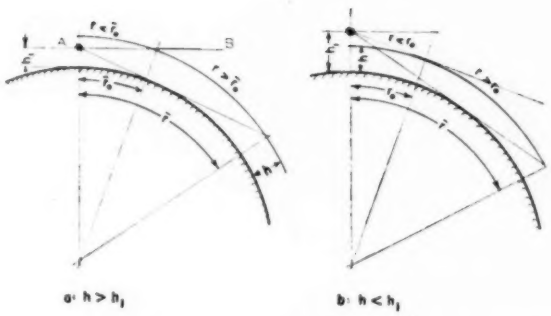


FIG. 2. h_1 = height of transmitter, h = height of receiver.

3. RAY THEORY

In this section we shall show briefly how one can derive from our solution given in (16) and (17) the geometric-optical approximation, which is valid for short wave-lengths inside the lit region. No detailed discussion of this subject will be attempted here, since this has been done adequately originally by van d. Pol and Bremmer and later by Burrows and Norton,[1] and since, furthermore, our main interest in this investigation lies in determining the field in the vicinity of the horizon. Now, according to Fermat's principle, a ray is a curve connecting the transmitter and receiver such that along it the travel-time is stationary. Such a curve is characterized by a definite angle θ which it makes with the vertical at the transmitter, and this value of θ is associated with a particular value of k ($<k_0$) in the integrals (16) and (17), through the relation[7] $k=k_0\mu(h_1)\sin\theta$. Hence our task is to find the values of k in the path of integration for which the travel-time is stationary. This travel-time is equal to the limit approached by the phase of the wave as the frequency (or k_0) is increased indefinitely, because the *beginning* of a pulse starting from quiescence is controlled by the high frequency components. It follows that our procedure is first to obtain asymptotic expressions for the integrands valid for large wave numbers k_0, then find the values of k for which the phase is stationary (if such points exist), and finally to evaluate the integrals around the points of stationary phase.

To begin with, the integrals in (16) and (17) can, under certain conditions[8] which are met in our case, be transformed by the use of the relation

$$\int_0^\infty J_0(kr)\psi(k)kdk=\tfrac{1}{2}\int_{-i\infty}^{\infty}H_0^{(2)}(kr)\psi(k)kdk. \quad (29)$$

In the limit of large r (in comparison with λ) we have

$$H_0^{(2)}(kr)\rightarrow(2/\pi kr)^{\frac{1}{2}}\exp[i\pi/4-ikr], \quad (30)$$

and (29) reduces to

$$\int_0^\infty J_0(kr)\psi(k)kdk \rightarrow[e^{i\pi/4}/(2\pi r)^{\frac{1}{2}}]\int_{-i\infty}^{\infty}e^{-ikr}\psi(k)k^{\frac{1}{2}}dk. \quad (31)$$

Taking now the integral in (16) we write it in the form

$$\Psi_1\simeq\frac{\pi\exp[i\omega t-i\pi/4]}{3(qk_0^2)^{\frac{1}{3}}(2\pi r)^{\frac{1}{2}}} \times\int_{-i\infty}^{\infty}e^{-ikr}[V(k)+W(k)]k^{\frac{1}{2}}dk, \quad h>h_1, \quad (32)$$

$$V(k)\equiv(u_1u)^{\frac{1}{2}}H_{1/3}^2(2u^{\frac{3}{2}}/3)H_{1/3}^{(1)}(2u_1^{\frac{3}{2}}/3), \quad (33)$$

$$W(k)\equiv(u_1u)^{\frac{1}{2}}(B/A)H_{1/3}^{(2)}(2u^{\frac{3}{2}}/3)H_{1/3}^{(2)}(2u_1^{\frac{3}{2}}/3).$$

For large wave numbers k_0, and as long as k is less than $k_0(1+qh_1)^{\frac{1}{3}}$ (and therefore also less than $k_0(1+qh)^{\frac{1}{3}}$), both u_1 and u are large, so that we obtain from (24)

$$e^{-ikr}V(k)\rightarrow(3/\pi)(uu_1)^{-\frac{1}{4}}\exp[-ip_1(k)], \quad p_1(k)=kr+(2u^{\frac{3}{2}}/3)-(2u_1^{\frac{3}{2}}/3). \quad (34)$$

The point k of stationary phase is determined from

$$\partial p_1(k)/\partial k=r-(2x/q)[(1+qh-x^2)^{\frac{1}{2}} -(1+qh_1-x^2)^{\frac{1}{2}}]=0, \quad (35)$$
$$x\equiv(k/k_0)=(1+qh_1)^{\frac{1}{3}}\sin\theta.$$

As x varies from *zero* to its maximum possible value of $(1+qh_1)^{\frac{1}{3}}$, r changes from *zero* to $\bar{r}_0$, where

$$q\bar{r}_0=2[1+qh_1]^{\frac{1}{2}}\cdot[q(h-h_1)]^{\frac{1}{2}} \quad (36)$$

defines the ray which starts out horizontally at the source. In real space the rays defined by (35) are the direct rays emitted above the transmitter's horizon.[9]

[8] Reference 5, pp. 305 to 307.

[9] Above line AB in Fig. 2.

If the receiver is located above h_1 but below the transmitter's horizon ($r>\bar{r}_0$), the phase of the term $e^{-ikr}V(k)$ in (32) does not have a stationary point, so that this term's contribution becomes small relative to the second term's, on account of its highly oscillatory nature. The direct ray reaching the receiver in this case arises from a stationary point k of $e^{-ikr}W(k)$, where $k_0<k<k_0(1+qh_1)^{\frac{1}{2}}$. This term has also another stationary point at $k<k_0$, which gives the reflected ray for any position of the receiver in the lit zone. When $k_0<k<k_0(1+qh_1)^{\frac{1}{2}}$ we put

$$w=(2/3)(k_0/q)[(k^2/k_0^2)-1]^{\frac{3}{2}}, \quad (2u_0^{\frac{1}{2}}/3)=e^{-i3\pi/2}w, \tag{37}$$

and obtain from (20)

$$\frac{B}{A}=\left\{\frac{\beta[I_{-1/3}(w)+e^{-i\pi/3}I_{1/3}(w)]+(3w/2)^{\frac{1}{2}}[I_{2/3}(w)+e^{-i\pi/3}I_{-2/3}(w)]}{\beta[I_{-1/3}(w)+e^{i\pi/3}I_{1/3}(w)]+(3w/2)^{\frac{1}{2}}[I_{2/3}(w)+e^{i\pi/3}I_{-2/3}(w)]}\right\}. \tag{38}$$

For large w

$$I_\nu(w)\rightarrow[e^w/(2\pi w)^{\frac{1}{2}}][1-\{(4\nu^2-1)/8w\}+\cdots], \quad B/A\rightarrow e^{-i\pi/3}, \tag{39}$$

and

$$e^{-ikr}W(k)\rightarrow(3/\pi)(uu_1)^{-\frac{1}{2}}\exp[(i\pi/2)-ip_2(k)], \quad p_2(k)=kr+(2u^{\frac{3}{2}}/3)+(2u_1^{\frac{3}{2}}/3), \tag{40}$$

$$\partial p_2(k)/\partial k=r-(2x/q)[(1+qh_1-x^2)^{\frac{1}{2}}+(1+qh-x^2)^{\frac{1}{2}}]=0. \tag{41}$$

As x varies from its minimum value of 1 to its maximum value of $(1+qh_1)^{\frac{1}{2}}$, r, as given by (41), changes from $\bar{r}$ to $\bar{r}_0$, where $\bar{r}$ denotes the horizontal distance from transmitter when the receiver is on the horizon, and is given by

$$q\bar{r}=2(qh_1)^{\frac{1}{2}}+2(qh)^{\frac{1}{2}}. \tag{42}$$

These are the direct rays emitted below the transmitter's horizon, thus complementing the direct rays not included in $p_1(k)$.

For $k<k_0$ we have, on using (25),

$$e^{-ikr}W(k)\rightarrow R(k)(3/\pi)(uu_1)^{-\frac{1}{2}}\exp[-ip_3(k)], \quad p_3(k)=kr-(4u_0^{\frac{3}{2}}/3)+(2u^{\frac{3}{2}}/3)+(2u_1^{\frac{3}{2}}/3), \tag{43}$$

$$\partial p_3(k)/\partial k=r-(2x/q)[(1+qh_1-x^2)^{\frac{1}{2}}+(1+qh-x^2)^{\frac{1}{2}}-2(1-x^2)^{\frac{1}{2}}]=0. \tag{44}$$

Here, as x varies from 0 to 1, r changes from *zero* to $\bar{r}$. These are the rays *reflected* at the surface of the earth, $R(k)$ representing the reflection coefficient for a plane earth, with the divergence factor arising from the integration around the stationary point, as is shown in the Appendix.

When the elevation of the receiver is less than that of the transmitter ($h<h_1$) we have to use Ψ_2 given in (17). Here the second term is identical with $W(k)$ in (32), and therefore leads to the phases $p_2(k)$ and $p_3(k)$, which are symmetrical in h_1 and h. In the case of $p_2(k)$

$$k_0<k<k_0(1+qh)^{\frac{1}{2}}, \quad q\bar{r}>qr>qr_0=2[1+qh]^{\frac{1}{2}}\cdot[q(h_1-h)]^{\frac{1}{2}}, \tag{45}$$

where r_0 denotes the horizontal distance when the transmitter is on the horizon of the receiver (see

TABLE I. Location of the points $x(=k^*/k_0)$ of stationary phase, and the rays associated with them.

	Phase	Eq.	Limits of x		Limits of r		Nature of ray	Sign of $\frac{\partial^2 p[k^*]}{\partial k^2}$
			x_1	x_2	$r(x_1)$	$r(x_2)$		
$h>h_1$	$p_1(k)$	35	0	$(1+qh_1)^{\frac{1}{2}}$	0	$\bar{r}_0$	Direct	−
	$p_2(k)$	41	1	$(1+qh_1)^{\frac{1}{2}}$	$\bar{r}$	$\bar{r}_0$	Direct	+
	$p_3(k)$	44	0	1	0	$\bar{r}$	Reflected	−
$h<h_1$	$p_4(k)$	47	0	$(1+qh)^{\frac{1}{2}}$	0	r_0	Direct	−
	$p_2(k)$	41	1	$(1+qh)^{\frac{1}{2}}$	$\bar{r}$	r_0	Direct	+
	$p_3(k)$	44	0	1	0	$\bar{r}$	Reflected	−

Fig. 2). When the receiver is closer to the transmitter ($r<r_0$), the direct ray arises from a stationary point of the first term in the integrand of (17):

$$e^{-ikr}(u_1u)^{\frac{1}{2}}H_{1/3}^{(2)}(2u_1^{\frac{3}{2}}/3)H_{1/3}^{(1)}(2u^{\frac{3}{2}}/3)\rightarrow(3/\pi)(uu_1)^{-\frac{1}{2}}\exp[-ip_4(k)]. \tag{46}$$

$$p_4(k)\equiv kr+(2u_1^{\frac{3}{2}}/3)-(2u^{\frac{3}{2}}/3),\quad \partial p_4(k)/\partial k=r-(2x/q)[(1+qh_1-x^2)^{\frac{1}{2}}-(1+qh-x^2)^{\frac{1}{2}}]=0. \tag{47}$$

As x varies from *zero* to $(1+qh)^{\frac{1}{2}}$, r changes from *zero* to r_0, thus complementing the rays arising from the stationary point in the second term. The reflected ray in the case $h<h_1$ arises from the second term, and is given by (43) and (44). The location of the points of stationary phase is summarized in Table I. Figure 2 illustrates the geometry of the problem.

We shall now indicate how one obtains the amplitudes of the rays by integrating around the points of stationary phase. It will suffice to illustrate the procedure in the case of the first term of Ψ_2 in (17), which yields the direct ray given in (47). We have first, on using (31) and then (46),

$$\begin{aligned}\Psi_2'\equiv-\frac{i\pi\exp[i\omega t]}{3(qk_0^2)^{\frac{1}{2}}}\int_0^\infty J_0(kr)kdk(u_1u)^{\frac{1}{2}}H_{1/3}^{(2)}(2u_1^{\frac{3}{2}}/3)H_{1/3}^{(1)}(2u^{\frac{3}{2}}/3)\\ \simeq\frac{\pi\exp[i\omega t-i\pi/4]}{3(qk_0^2)^{\frac{1}{2}}(2\pi r)^{\frac{1}{2}}}\int_{-i\infty}^{\infty}e^{-ikr}(u_1u)^{\frac{1}{2}}H_{1/3}^{(2)}(2u_1^{\frac{3}{2}}/3)H_{1/3}^{(1)}(2u^{\frac{3}{2}}/3)(k)^{\frac{1}{2}}dk\\ \simeq\frac{\exp[i\omega t-i\pi/4]}{(qk_0^2)^{\frac{1}{2}}(2\pi r)^{\frac{1}{2}}}\int_{-i\infty}^{\infty}(uu_1)^{-\frac{1}{2}}\exp[-ip_4(k)]k^{\frac{1}{2}}dk.\end{aligned} \tag{48}$$

Let the point k at which $p_4(k)$ is stationary be denoted by k^*; then we have in the neighborhood of k^*

$$p_4(k)=p_4(k^*)+\tfrac{1}{2}[\partial^2p_4(k^*)/\partial k^2]t^2+\cdots,\quad t=k-k^*. \tag{49}$$

Using this expansion in (48) we obtain

$$\begin{aligned}\int_{-i\infty}^{\infty}(uu_1)^{-\frac{1}{2}}\exp[-ip_4(k)]k^{\frac{1}{2}}dk\simeq\exp[-ip_4(k^*)][(uu_1)^{-\frac{1}{2}}k^{\frac{1}{2}}]_{k=k^*}\cdot\int_{-\infty}^{\infty}\exp[-(i\partial^2p_4/2\partial k^2)t^2]dt\\ =(k^*)^{\frac{1}{2}}(q/k_0)^{\frac{1}{2}}(1+qh_1-x^2)^{-\frac{1}{2}}(1+qh-x^2)^{-\frac{1}{2}}\exp[i\pi/4-ip_4(k^*)]\left|\frac{2\pi}{\partial^2p_4/\partial k^2}\right|^{\frac{1}{2}},\end{aligned} \tag{50}$$

where we have written x for k^*/k_0. Hence, the direct ray is given by

$$\Psi_2'\simeq\frac{\exp[i\omega t-ip_4(k^*)]x^{\frac{1}{2}}}{[k_0r\,|\,\partial^2p_4/\partial k^2\,|\,]^{\frac{1}{2}}(1+qh_1-x^2)^{\frac{1}{2}}(1+qh-x^2)^{\frac{1}{2}}}. \tag{51}$$

This can now be expressed explicitly in terms of r, h_1, and h by using (47). It follows from the latter that

$$x^2=[r^2/4R^2]\cdot[a+b+(4ab-q^2r^2)^{\frac{1}{2}}],\quad a\equiv1+qh_1,\quad b\equiv1+qh,\quad R\equiv[r^2+(h-h_1)^2]^{\frac{1}{2}}, \tag{52}$$

$$\partial^2p_4/\partial k^2=-[r(4ab-q^2r^2)^{\frac{1}{2}}/2k_0x(a-x^2)^{\frac{1}{2}}(b-x^2)^{\frac{1}{2}}],$$

$$\Psi_2'\simeq\frac{\exp[i\omega t-ip_4(k^*)]x2^{\frac{1}{2}}}{r(4ab-q^2r^2)^{\frac{1}{2}}}=\frac{\exp[i\omega t-ip_4(k^*)][a+b+(4ab-q^2r^2)^{\frac{1}{2}}]^{\frac{1}{2}}}{2^{\frac{1}{2}}R(4ab-q^2r^2)^{\frac{1}{2}}}, \tag{53}$$

$$p_4(k^*)=[k_0/(3\cdot2^{\frac{1}{2}}R)]\cdot[(a+b+qR)^{\frac{1}{2}}+(a+b-qR)^{\frac{1}{2}}]\cdot[q^2r^2+2a^2+2b^2-(a+b)(4ab-q^2r^2)^{\frac{1}{2}}]. \tag{54}$$

The amplitude of the direct ray reduces to R^{-1} if powers of q higher than the first are neglected.

As to the normal mode theory, which is useful in the shadow zone, we shall merely mention here that it can be obtained by evaluating the integral in (32) at the residues of (B/A) defined in (20). For details of this calculation in the case of strong absorption ($\beta\gg1$) the reader is referred to reference 5, pages 307 and 308.

TABLE II.

	$k<k_0$	$k>k_0$
w_0	$(3v/2)^{\frac{1}{2}}$	$e^{-i\pi}(3w/2)^{\frac{1}{2}}$
w	$z_2+(3v/2)^{\frac{1}{2}}$	$z_2-(3w/2)^{\frac{1}{2}}$
w_1	$z_1+(3v/2)^{\frac{1}{2}}$	$z_1-(3w/2)^{\frac{1}{2}}$
$v_0(19)$	$[L^2(\epsilon_1-1)/4H^2\epsilon_1^{\frac{1}{2}}]+(3v/\epsilon_12)^{\frac{1}{2}}$	$[L^2(\epsilon_1-1)/4H^2\epsilon_1^{\frac{1}{2}}]-(3w/\epsilon_12)^{\frac{1}{2}}$
(k/k_0)	$1-\frac{H^2}{2L^2}(12v)^{\frac{1}{2}}-\frac{H^4}{8L^4}(12v)^{4/3}+\cdots$	$1+\frac{H^2}{2L^2}(12w)^{\frac{1}{2}}-\frac{H^4}{8L^4}(12w)^{4/3}+\cdots$
$p_2(k)$	$p_2(k_0)-pv^{\frac{1}{2}}+[x/4(z_1z_2)^{\frac{1}{2}}](3v/2)^{4/3}+\cdots$	$p_2(k_0)+pw^{\frac{1}{2}}+[x/4(z_1z_2)^{\frac{1}{2}}](3v/2)^{4/3}+\cdots$

4. THE FIELD IN THE VICINITY OF THE HORIZON IN CASE THE TRANSMITTER AND RECEIVER ARE ELEVATED SEVERAL NATURAL UNITS OF HEIGHT ABOVE THE GROUND

We now turn to our principal task, which is the determination of the field in the vicinity of the horizon. It will be noted from Table I that as $r\to\bar{r}$ (the horizon) the stationary point of $p_2(k)$ approaches k_0 from above, while the stationary point of $p_3(k)$ approaches the same limit from below. This corresponds to the coalescence of the direct ray with the reflected ray on the horizon. Since both of these stationary points arise from the second terms (containing the B/A factor) in the integrands of (16) and (17), and since the principal contribution to the integrals occurs near the stationary points, it follows that, except for the case when the transmitter or receiver are on the ground, the field in the neighborhood of the horizon can be determined approximately by integrating the second terms in the integrands of (16) and (17) over a small range of k on either side of k_0. When the transmitter or receiver, or both, are on the ground, both terms in the integrands of (16) or (17) need to be considered.

In the sequel it will be convenient to define a natural unit of height, H, and a natural unit of horizontal distance, L, through

$$H=(k_0^2q)^{-\frac{1}{3}},\quad L=2(kq^2)^{-\frac{1}{3}},\quad z_1=h_1/H,\quad z_2=h/H,\quad x=r/L,\quad p=(3/2)^{\frac{1}{2}}(x-\bar{x}). \tag{55}$$

If λ is expressed in centimeters then, for a mean radius of the earth of 6371 km and $a_e=(4/3)a$, we have

$$H(m)=2.208\lambda^{\frac{2}{3}},\quad L(\text{km})=6.124\lambda^{\frac{1}{3}}. \tag{56}$$

Further, we write

$$(k^2/k_0^2)-1=(3qw/2k_0)^{\frac{1}{2}},\quad k>k_0;\qquad 1-(k^2/k_0^2)=(3vq/2k_0)^{\frac{1}{2}},\quad k<k_0, \tag{57}$$

and tabulate the expansions of some relevant quantities in Table II. (B/A) is given in terms of w for $k>k_0$ in (38), while for $k<k_0$ we have

$$\frac{B}{A}=\left\{\frac{\beta[J_{-1/3}(v)-e^{-i\pi/3}J_{1/3}(v)]+(3v/2)^{\frac{1}{2}}[J_{2/3}(v)+e^{-i\pi/3}J_{-2/3}(v)]}{\beta[J_{-1/3}(v)-e^{i\pi/3}J_{1/3}(v)]+(3v/2)^{\frac{1}{2}}[J_{2/3}(v)+e^{i\pi/3}J_{-2/3}(v)]}\right\}. \tag{58}$$

We shall limit the discussion to the case of strong absorption when β is large, so that

$$\frac{B}{A}\longrightarrow-\left[\frac{H_{1/3}^{(1)}(2u_0^{\frac{3}{2}}/3)}{H_{1/3}^{(2)}(2u_0^{\frac{3}{2}}/3)}\right]=\frac{[J_{-1/3}(v)-e^{-i\pi/3}J_{1/3}(v)]}{[J_{-1/3}(v)-e^{i\pi/3}J_{1/3}(v)]}=\frac{[I_{-1/3}(w)+e^{-i\pi/3}I_{1/3}(w)]}{[I_{-1/3}(w)+e^{i\pi/3}I_{1/3}(w)]}. \tag{59}$$

Now

$$\beta_v=(iv_0^{\frac{1}{2}}/\epsilon_1^{\frac{1}{2}})\sim i[(k_0/q)^{\frac{1}{2}}]\cdot[(\epsilon_1-1)^{\frac{1}{2}}/\epsilon_1];\quad \beta_h=\beta_v\epsilon_1, \tag{60}$$

where the subscripts v and h signify vertical and horizontal polarization, respectively. Table III shows in four typical cases of ground constants that for wave-lengths less than a meter $|\beta_v|\gg1$, and that this is true at all frequencies for $|\beta_h|$. Relation (59) is strictly true when the ground is a perfect reflector,[8] i.e., when the boundary condition at the ground is $\Psi=0$.

It follows from the preceding discussion that in the vicinity of the horizon

$$\psi \simeq \frac{i\pi \exp[i\omega t]}{3(qk_0^2)^{\frac{1}{3}}} \int_0^\infty J_0(kr)kdk(u_1u)^{\frac{1}{2}}H_{1/3}^{(2)}(2u^{\frac{3}{2}}/3)H_{1/3}^{(2)}(2u_1^{\frac{3}{2}}/3)\frac{H_{1/3}^{(1)}(2u_0^{\frac{3}{2}}/3)}{H_{1/3}^{(2)}(2u_0^{\frac{3}{2}}/3)}$$

$$\simeq -\frac{\pi \exp[i\omega t - i\pi/4]}{3(qk_0^2)^{\frac{1}{3}}(2\pi r)^{\frac{1}{2}}} \int_{-i\infty}^{\infty} e^{-ikr}(u_1u)^{\frac{1}{2}}H_{1/3}^{(2)}(2u^{\frac{3}{2}}/3)H_{1/3}^{(2)}(2u_1^{\frac{3}{2}}/3)\frac{H_{1/3}^{(1)}(2u_0^{\frac{3}{2}}/3)}{H_{1/3}^{(2)}(2u_0^{\frac{3}{2}}/3)}k^{\frac{1}{2}}dk \tag{61}$$

$$\simeq \frac{\pi^{\frac{1}{2}} \exp[i\omega t - i\pi/4]}{3(12)^{\frac{1}{2}}(rL)^{\frac{1}{2}}} \left\{ \int_0^\infty e^{-ikr}(u_1u)^{\frac{1}{2}}H_{1/3}^{(2)}(2u^{\frac{3}{2}}/3)H_{1/3}^{(2)}(2u_1^{\frac{3}{2}}/3)\left[\frac{I_{-1/3}(w)+e^{-i\pi/3}I_{1/3}(w)}{I_{-1/3}(w)+e^{i\pi/3}I_{1/3}(w)}\right]\frac{dw}{w^{\frac{1}{2}}} \right.$$

$$\left. + \int_0^\infty e^{-ikr}(u_1u)^{\frac{1}{2}}H_{1/3}^{(2)}(2u^{\frac{3}{2}}/3)H_{1/3}^{(2)}(2u_1^{\frac{3}{2}}/3)\left[\frac{J_{-1/3}(v)-e^{-i\pi/3}J_{1/3}(v)}{J_{-1/3}(v)-e^{i\pi/3}J_{1/3}(v)}\right]\frac{dv}{v^{\frac{1}{2}}} \right\}, \tag{62}$$

where u, u_1, and k take on their respective expressions in terms of w and v, as shown in Table II. The only approximation made in going from (61) to (62) was in writing $k_0^{\frac{1}{2}}$ for $k^{\frac{1}{2}}$, which is justified because H^2/L^2 $(=1.3\times10^{-7}\lambda^{\frac{1}{2}}_{\text{cm}})$ is extremely small in the microwave region.

We shall evaluate the integral in (62) for the case when z_1 and z are several units. We can then write

$$e^{-ikr}(u_1u)^{\frac{1}{2}}H_{1/3}^{(2)}(2u^{\frac{3}{2}}/3)H_{1/3}^{(2)}(2u_1^{\frac{3}{2}}/3) \rightarrow (3/\pi)(uu_1)^{-\frac{1}{4}} \exp[i5\pi/6 - ip_2(k)], \tag{63}$$

$$\psi \simeq \frac{\exp[i7\pi/12 + i\omega t - ip_2(k_0)]}{(12)^{\frac{1}{2}}(\pi rL)^{\frac{1}{2}}(z_1z_2)^{\frac{1}{4}}} \left\{ \int_0^\infty \exp[-ipw^{2/3} - isw^{4/3}]\left(e^{-i\pi/3} + \frac{e^{i\pi/3}[I_{-1/3}(w)-I_{1/3}(w)]}{[I_{-1/3}(w)+e^{i\pi/3}I_{1/3}(w)]}\right)\frac{dw}{w^{\frac{1}{2}}} \right.$$

$$\left. + \int_0^\infty \exp[ipv^{2/3} - isv^{4/3}]\left(\frac{J_{-1/3}(v)-e^{-i\pi/3}J_{1/3}(v)}{J_{-1/3}(v)-e^{i\pi/3}J_{1/3}(v)}\right)\frac{dv}{v^{\frac{1}{2}}} \right\}, \tag{64}$$

$$\int_0^\infty \exp[-ipw^{2/3} - isw^{4/3}](dw/w^{\frac{1}{2}}) = [(12)^{\frac{1}{2}}/\Omega]\int_0^\infty \exp[-i(2\tau^{\frac{1}{2}}t + t^2)]dt, \tag{65}$$

$$\Omega^2 = [(1/z_1^{\frac{1}{2}}) + (1/z_2^{\frac{1}{2}})], \quad \tau = (x-\bar{x})^2/\Omega^2, \quad s = (3/2)^{4/3}\bar{x}/4(z_1z_2)^{\frac{1}{2}} = (3/2)^{4/3}\Omega^2/4;$$

$$\int_0^\infty \exp[ipv^{\frac{3}{2}}]\left[\frac{J_{-1/3}(v)-e^{-i\pi/3}J_{1/3}(v)}{J_{-1/3}(v)-e^{i\pi/3}J_{1/3}(v)}\right]\frac{dv}{v^{\frac{1}{2}}}$$

$$= e^{i\pi/3}\int_0^\infty \exp\{[(-3^{\frac{1}{2}}/2) + (i/2)]px^{\frac{3}{2}}\}\left[\frac{I_{-1/3}(x)-I_{1/3}(x)}{I_{-1/3}(x)+e^{-i\pi/3}I_{1/3}(x)}\right], \tag{66}$$

$$\psi = \frac{\exp[i\omega t - ip_2(k_0)]}{2(r\bar{r})^{\frac{1}{2}}} \left\{ \frac{2e^{i\pi/4}}{\pi^{\frac{1}{2}}} \int_0^\infty \exp[-i(2\tau^{\frac{1}{2}}t + t^2)]dt - \frac{(2/3)^{\frac{1}{2}}\Omega e^{-i\pi/12}}{\pi^{\frac{1}{2}}}F(p) \right\}, \tag{67}$$

TABLE III. Values of $|\beta_v|$ and $|\beta_h|$ for four typical ground constants.

λ	1 cm	10 cm	100 cm	10 m	600 m	
$\|\beta_v\|$	601	279	129	49.8	1.87	
$\|\beta_h\|$	2400	1120	523	359	672	$\epsilon=4$, $\sigma=10^{-13}$ e.m.u.
$\|\beta_v\|$	436	202	93.9	43.3	3.40	Dry soil
$\|\beta_h\|$	3920	1820	845	397	368	$\epsilon=9$, $\sigma=3\times10^{-14}$ e.m.u.
$\|\beta_v\|$	346	161	74.5	34.3	2.64	Average soil
$\|\beta_h\|$	5190	2410	1120	525	476	$\epsilon=15$, $\sigma=5\times10^{-14}$ e.m.u.
$\|\beta_v\|$	154	69.3	17.0	2.53	0.084	Sea water
$\|\beta_h\|$	12300	5920	5260	7590	15020	$\epsilon=80$, $\sigma=5\times10^{-11}$ e.m.u.

TABLE IV. Values of $F(p)$ and $G(p)$ defined in (68) and (78), respectively; R and I denote real and imaginary parts.

p	−4.0	−3.5	−3.0	−2.5	−2.0	−1.5	−1.0	−0.5	0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
$RF(p)$	−2.35	.18	1.70	2.21	2.21	2.04	1.84	1.64	1.44	1.23	1.03	.84	.68	.55	.45	.38	.33
$IF(p)$	−2.79	−3.15	−2.18	−1.10	−.36	.02	.14	.10	0	−.11	−.20	−.26	−.27	−.26	−.24	−.21	−.19
				$RG(p)$	8.38	5.96	3.62	2.49	1.84	1.24	.67	.27	.07				
				$IG(p)$	−1.63	2.89	3.24	2.22	1.06	.23	−.16	−.23	−.16				

$$F(p)=\int_0^\infty \exp[-ipx^{\frac{1}{2}}]\frac{dx}{x^{\frac{1}{2}}}\left[\frac{I_{-1/3}(x)-I_{1/3}(x)}{I_{-1/3}(x)+e^{i\pi/3}I_{1/3}(x)}\right]$$
$$+\int_0^\infty \exp\{[(-3^{\frac{1}{2}}/2)+(i/2)]px^{\frac{1}{3}}\}\frac{dx}{x^{\frac{1}{2}}}\left[\frac{I_{-1/3}(x)-I_{1/3}(x)}{I_{-1/3}(x)+e^{-i\pi/3}I_{1/3}(x)}\right]. \quad (68)$$

In deriving (67) we have neglected the contribution from the $v^{4/3}$ and $w^{4/3}$ terms in the exponents of the *second* and *third* terms in the integrand of (64). This is justified if the principal contribution to $F(p)$ arises from a small range in x near the origin, or if $s\ll 1$. In any case, the error involved can be estimated by using instead of $F(p)$ in (67) the quantity $[F(p)+is(d^2F/dp^2)]$. If $s(d^2F/dp^2)$ is small, it can be added to $F(p)$ to improve the accuracy; if it is not small, (67) cannot be relied upon.

Values of $F(p)$ at intervals of 0.5 obtained from (68) are shown[10] in Table IV. From the second differences, approximate values of d^2F/dp^2 were obtained, and are plotted in Fig. 3.

The integral in (67) can be expressed in terms of the tabulated Fresnel[11] integrals, as follows:

$$(2/\pi^{\frac{1}{2}})(e^{i\pi/4})\int_0^\infty \exp[-i(2\tau^{\frac{1}{2}}t+t^2)]dt=e^{i\tau}\{1-(1+i)[C(\tau)-iS(\tau)]\},\quad \tau^{\frac{1}{2}}>0,$$
$$=e^{i\tau}\{1+(1+i)[C(\tau)-iS(\tau)]\},\quad \tau^{\frac{1}{2}}<0, \quad (69)$$

where $\tau^{\frac{1}{2}}[=(x-\bar{x})/\Omega]$ is positive for $r>\bar{r}$, and negative for $r<\bar{r}$. For large positive p (in the shadow zone)

$$F(p)\to(3/2p)e^{-i\pi/6},\quad \int_0^\infty \exp[-i(2\tau^{\frac{1}{2}}t+t^2)]dt\to -i/2\tau^{\frac{1}{2}}, \quad (70)$$

and these leading terms just cancel each other in (67).

5. THE FIELD IN THE VICINITY OF THE HORIZON WHEN THE TRANSMITTER IS ON THE GROUND AND THE RECEIVER IS ELEVATED SEVERAL NATURAL UNITS OF HEIGHT ABOVE THE GROUND

When $h_1=0$, but $h>0$, we have $u_1=u_0$, and Eqs. (16), (15), and (20) yield

$$\Psi=[2e^{i\omega t}/(qk_0^2)^{\frac{1}{2}}]\int_0^\infty J_0(kr)kdku^{\frac{1}{2}}H_{1/3}^{(2)}(2u^{\frac{1}{2}}/3)/Q(k), \quad (71)$$

$$Q(k)\equiv[\beta u_0^{\frac{1}{2}}H_{1/3}^{(2)}(2u_0^{\frac{1}{2}}/3)-u_0H_{-2/3}^{(2)}(2u_0^{\frac{1}{2}}/3)]. \quad (72)$$

Proceeding now as in the previous section we obtain, in the case of strong absorption and a value of z_2 of several units,

$$\Psi\simeq\frac{2^{\frac{1}{2}}\exp[i\omega t+i\pi/4]}{(qk_0^2)^{\frac{1}{2}}(\pi r)^{\frac{1}{2}}}\int_{-i\infty}^\infty e^{-ikr}k^{\frac{1}{2}}dku^{\frac{1}{2}}H_{1/3}^{(2)}(2u^{\frac{1}{2}}/3)/Q(k), \quad (73)$$

[10] We are indebted to the Mathematical Tables Project for the use of their manuscript of Bessel functions of order $\frac{1}{3}$.
[11] Jahnke-Emde, *Tables of Functions* (Dover Publications, New York, 1933), p. 35.

$$Q \simeq \beta u_0^{\frac{1}{2}} H_{1/3}^{(2)}(2u^{\frac{1}{2}}/3) = [(2/3^{\frac{1}{2}}) \cdot (3v/2)^{\frac{1}{2}}] \cdot [iJ_{-1/3}(v) + e^{-i\pi/6} J_{1/3}(v)]\beta, \quad k < k_0,$$

$$= [(2/3^{\frac{1}{2}}) \cdot (3w/2)^{\frac{1}{2}}] \cdot [iI_{-1/3}(w) - e^{-i\pi/6} I_{1/3}(w)]\beta, \quad k > k_0, \tag{74}$$

$$e^{-ikr} u^{\frac{1}{2}} H_{1/3}^{(2)}(2u^{\frac{1}{2}}/3) \rightarrow (3/\pi^{\frac{1}{2}}) u^{-\frac{1}{4}} \exp[i5\pi/12 - ip_5(k)], \quad p_5(k) = kr + (2u^{\frac{1}{2}}/3), \tag{75}$$

$$p_5(k) = p_5(k_0) - pv^{\frac{1}{2}} + s'v^{4/3} + \cdots = p_5(k_0) + pw^{\frac{1}{2}} + s'w^{4/3} + \cdots, \quad s' = (3/2)^{4/3}/4z_2^{\frac{1}{2}}, \tag{76}$$

$$\Psi \simeq \frac{\exp\{i[\omega t - (\pi/3) - k_0 r - (2z_2^{\frac{1}{2}}/3)]\}(3\lambda/2\pi a_e)^{\frac{1}{2}}}{\pi(r\bar{r})^{\frac{1}{2}}\gamma(\epsilon_1 - 1)^{\frac{1}{2}}} G(p), \tag{77}$$

$$G(p) = \int_0^\infty \frac{\exp[-ipw^{\frac{1}{2}}]dw}{w^{\frac{1}{2}}[I_{-1/3}(w) + e^{i\pi/3} I_{1/3}(w)]} + \int_0^\infty \frac{\exp[ipv^{\frac{1}{2}}]dv}{v^{\frac{1}{2}}[J_{-1/3}(v) - e^{i\pi/3} J_{1/3}(v)]}$$

$$= \int_0^\infty \frac{\exp[-ipx^{\frac{1}{2}}]dx}{x^{\frac{1}{2}}[I_{-1/3}(x) + e^{i\pi/3} I_{1/3}(x)]} + e^{i\pi/3} \int_0^\infty \frac{\exp\{[(-3^{\frac{1}{2}}/2) + (i/2)]px^{\frac{1}{2}}\}dx}{x^{\frac{1}{2}}[I_{-1/3}(x) + e^{-i\pi/3} I_{1/3}(x)]}. \tag{78}$$

As before, the contribution from the s' terms in the exponents can be estimated by using instead of $G(p)$, $[G(p) + is'(d^2G/dp^2)]$.

Values of $G(p)$ are given in Table IV, and $|G(p)|$ is plotted in Fig. 4. Figure 3 gives approximate values of d^2G/dp^2, which may be used for the purpose of estimating the correction due to the s' term.

6. THE FIELD WHEN BOTH TRANSMITTER AND RECEIVER ARE ON THE GROUND (SURFACE WAVE)

The expression for the potential is obtained in this case by putting $u = u_0$ in (72):

$$\Psi = [2e^{i\omega t}/(qk_0^2)^{\frac{1}{2}}] \int_0^\infty J_0(kr) k dk u_0^{\frac{1}{2}} H_{1/3}^{(2)}(2u^{\frac{1}{2}}/3)/Q(k). \tag{79}$$

Making use of the identity

$$\frac{u_0^{\frac{1}{2}} H_{1/3}^{(2)}(2u_0^{\frac{1}{2}}/3)}{Q} = \frac{(qk_0^2)^{\frac{1}{2}}}{[(k_0^2 - k^2)^{\frac{1}{2}} + \gamma(k_1^2 - k^2)^{\frac{1}{2}}]} \cdot \{-i + [u_0 H_{1/3}^{(2)}(2u^{\frac{1}{2}}/3) - iu_0 H_{-2/3}^{(2)}(2u_0^{\frac{1}{2}}/3)]/Q\}, \tag{80}$$

we can write

$$\Psi = \phi + \psi, \quad \phi = -2ie^{i\omega t} \int_0^\infty \frac{J_0(kr)kdk}{[(k_0^2 - k^2)^{\frac{1}{2}} + \gamma(k_1^2 - k^2)^{\frac{1}{2}}]}, \tag{81}$$

$$\psi = 2e^{i\omega t} \int_0^\infty \frac{J_0(kr)kdk[u_0 H_{1/3}^{(2)}(2u_0^{\frac{1}{2}}/3) - iu_0 H_{-2/3}^{(2)}(2u_0^{\frac{1}{2}}/3)]}{[(k_0^2 - k^2)^{\frac{1}{2}} + \gamma(k_1^2 - k^2)^{\frac{1}{2}}]Q(k)}. \tag{82}$$

Here ϕ is Sommerfeld's potential of the surface wave for a *flat* earth, as can be seen by putting $h_1 = h = 0$ in (28). ψ is an *additive* potential which represents the effect of the curvature of the earth; it vanishes in the limit of an infinite radius.

Using the same methods as in the preceding two sections we obtain, for the case of strong absorption,

$$\Psi \simeq \frac{\exp[i(3\pi/4) - irk_0]q}{(2\pi rk_0)^{\frac{1}{2}}(\epsilon_1 - 1)\gamma^2} H(p'), \quad p' = (3/2)^{\frac{1}{2}}(r/L), \tag{83}$$

$$H(p') = \int_0^\infty \exp[-ip'x^{\frac{1}{2}}] \cdot [E(x) - iF(x)]dx + \int_0^\infty \exp\{p'[(-3^{\frac{1}{2}}/2) + (i/2)]x^{\frac{1}{2}}\} \cdot [E(x) + iF(x)]dx, \tag{84}$$

$$E(x) - iF(x) = \frac{e^{-i\pi/6}[I_{-2/3}(x) - I_{1/3}(x)] + i[I_{-1/3}(x) - I_{2/3}(x)]}{[I_{-1/3}(x) + e^{i\pi/3} I_{1/3}(x)]}. \tag{85}$$

In the microwave region we may use for ϕ the leading term in the asymptotic expansion

$$\phi \rightarrow -\frac{2i \exp[-ik_0 r]}{\gamma^2(\epsilon_1-1)k_0 r^2}\left[1+\frac{i(3\epsilon_1+2)}{k_0 r}+\cdots\right], \tag{86}$$

which follows by partial integration of van d. Pol's integral for ϕ.[12] Combining (83) with (86) we obtain for the surface wave in a spherical earth

$$\Psi \simeq \frac{\exp[-i(\pi/2)+ik_0 r]}{\gamma^2(\epsilon_1-1)k_0 r^2} g(p'), \quad g(p')=1-\frac{2e^{i\pi/4}}{3\pi^{\frac{1}{2}}}(p')^{\frac{1}{2}}H(p'). \tag{87}$$

The function $H(p')$ has been computed from (84), and $|g(p')|$ is shown by the dashed curve in Fig. 5, where it is compared with a corresponding function $G(\eta)$ ($\eta=0.521p'$) computed by van d. Pol and Bremmer,[13] who used the normal mode theory.

7. DISCUSSION OF RESULTS

In this section we shall apply our theory to several cases which have been computed exactly by van d. Pol and Bremmer,[1] using the normal mode theory. The theory developed here is best suited for computing the field on the horizon and its immediate proximity. It becomes increasingly inaccurate as one recedes either far into the lit zone or far into the shadow zone. In addition, we also restricted ourselves to the case of strong absorption through the assumption that the quantity β given in (21), (22), and (60) is large compared with unity. In the case of an elevated transmitter or receiver we have also assumed that the elevation is several units of natural height (z_1 or $z_2 \gg 1$), and that the quantities s, or s' defined in (66) and (76) are small. All these assumptions become valid as the wave-length is decreased.

Figures 6 and 7 show a comparison of our results with exact values obtained by van d. Pol and Bremmer for an elevated transmitter and receiver. The case of a grounded transmitter is

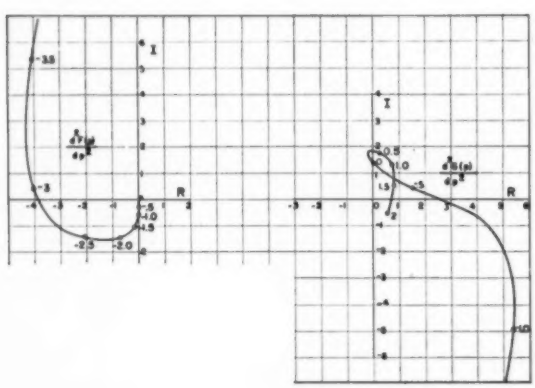


FIG. 3. Approximate values of $[d^2F(p)/dp^2]$ and $[d^2G(p)/dp^2]$. The numbers on the curves denote values of p.

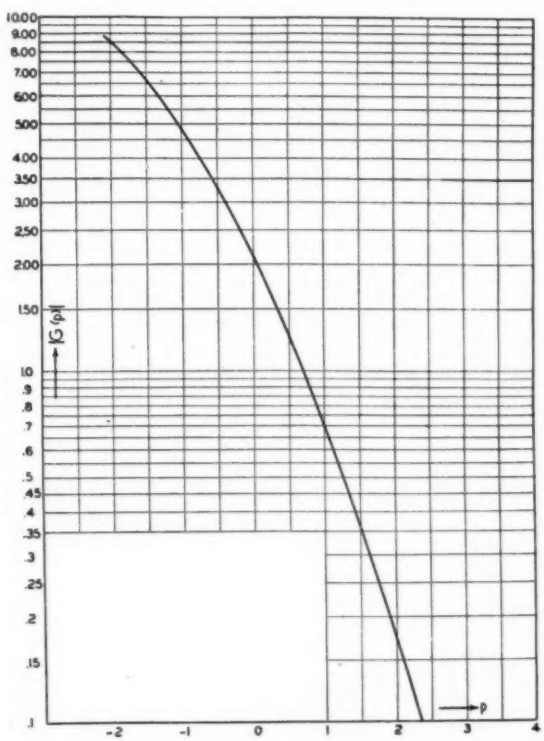


FIG. 4. The modulus of $G(p)$ defined in Eq. (78) and tabulated in Table IV.

[12] B. van d. Pol, Zeits. f. Hochfreqenztech. **37**, 152 (1931).
[13] B. van d. Pol and H. Bremmer, Phil. Mag. **24**, 173 (1937).

shown in Fig. 8. Some relevant quantities are tabulated in Table V, where the actual radius of the earth was used so as to conform to the calculations of van d. Pol and Bremmer. It is seen that in the cases treated the agreement near the horizon is sufficient for practical purposes for microwaves. Good agreement is also shown by Fig. 5 for the surface wave. Here our results do not go beyond those obtained by van d. Pol and Bremmer.

In a subsequent communication it will be shown that when the transmitter's height z_1 is small, equations (B) and (C) are to be multiplied by a factor $|(1+\beta z_1|$. When both z_1 and z_2 are small the potential $|\Psi|$ in (D) is to be multiplied by $|(1+\beta z_1)(1+\beta z_2)|$. Burrows and Gray[1] have computed a function $F(L)=(3^{\frac{1}{2}}/2\pi)\times|G[(8/9)^{\frac{1}{2}}p]|$, using the normal mode theory. These authors also give $F(L)$ for the case of a perfectly conducting earth.

APPENDIX

The Amplitude of the Reflected Wave

If we integrate the second term in (32) around the stationary point k^* of $p_3(k)$ defined in (44) we obtain, in a manner similar to the derivation of (51), for the amplitude Ψ_1' of the reflected wave

$$\Psi_1'=\frac{\exp[i\omega t-ip_3(k^*)]R(k^*)(qx)^{\frac{1}{2}}}{(rP)^{\frac{1}{2}}(1+qh_1-x^2)^{\frac{1}{4}}\cdot(1+qh_2-x^2)^{\frac{1}{4}}}. \qquad (a)$$

$$P\equiv 2(1+qh_1-x^2)^{\frac{1}{2}}-[2x^2/(1+qh_1-x^2)^{\frac{1}{2}}]+2(1+qh_2-x^2)^{\frac{1}{2}}-[2x^2/(1+qh_2-x^2)^{\frac{1}{2}}]-4(1-x^2)^{\frac{1}{2}}+[4x^2/(1-x^2)^{\frac{1}{2}}];\quad x=k^*/k_0. \qquad (b)$$

Here $R(k)$, the reflection coefficient for a plane earth, is given in (26). In deriving it we used the asymptotic expression (25), which is valid only if u_0 is large. This is

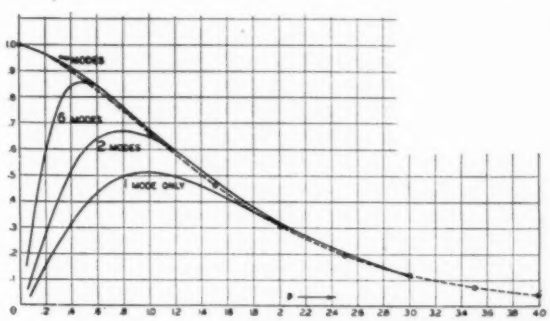


FIG. 5. Comparison of our results for the surface wave with those of van d. Pol and Bremmer (Phil. Mag. **24**, 173, 1937). – – – – $|g(p')|$, Eq. (87); ——— the corresponding function from normal mode theory using various modes.

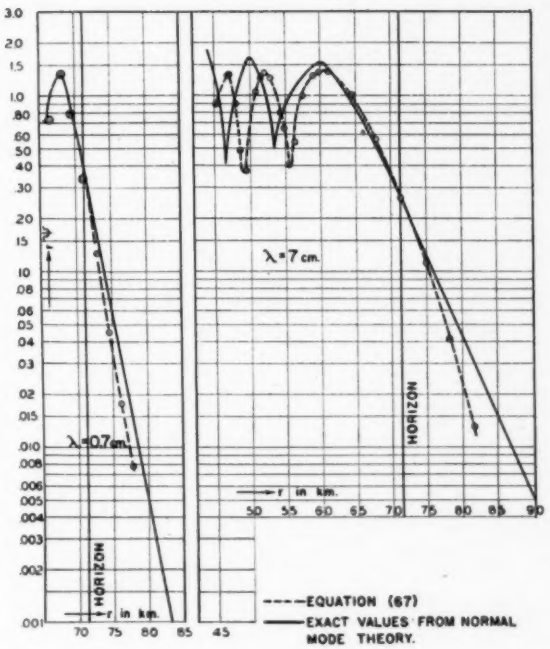


FIG. 6. Comparison of our results with those of van d. Pol and Bremmer.[2] Transmitter and receiver at a height of 100 m. $\epsilon=4$, $\sigma=10^{-13}$ e.m.u.

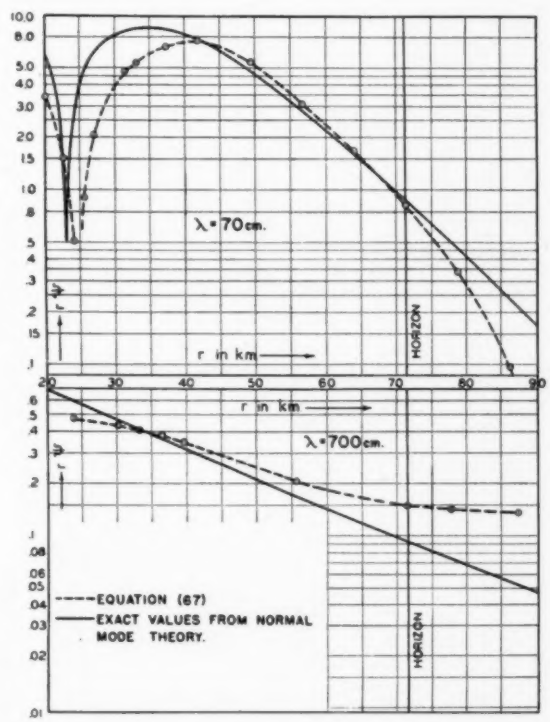


FIG. 7. Comparison of our results with those of van d. Pol and Bremmer.[2] Transmitter and receiver at a height of 100 m. $\epsilon=4$, $\sigma=10^{-13}$ e.m.u.

not true near grazing incidence, where u_0 becomes small. For microwaves, however, the spherical reflection coefficient is still well represented by $R(k)$ even at grazing incidence, except in the limiting case of a perfectly conducting ground.[14]

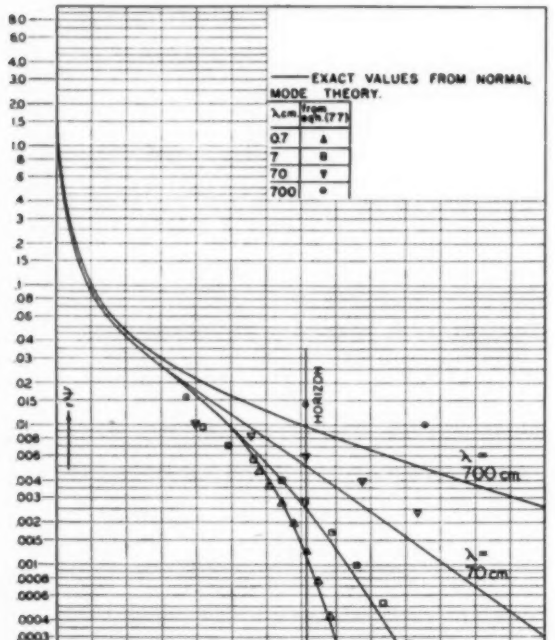


FIG. 8. Comparison of our results with those of van d. Pol and Bremmer.[1] Transmitter on the ground. Receiver at a height of 100 m. $\epsilon=4$, $\sigma=10^{-12}$ e.m.u.

TABLE V. Some quantities pertaining to the cases treated by van d. Pol and Bremmer;[1] $\epsilon=4$, $\sigma=10^{-13}$ e.m.u.

λ	s_1 or s_2	s	s'	$\lvert\beta_r\rvert$
0.7 cm	63	.11	.054	614
7 cm	14	.23	.12	285
70 cm	2.9	.50	.25	132
700 cm	.63	1.08	.54	56

[14] Reference 2, p. 268.

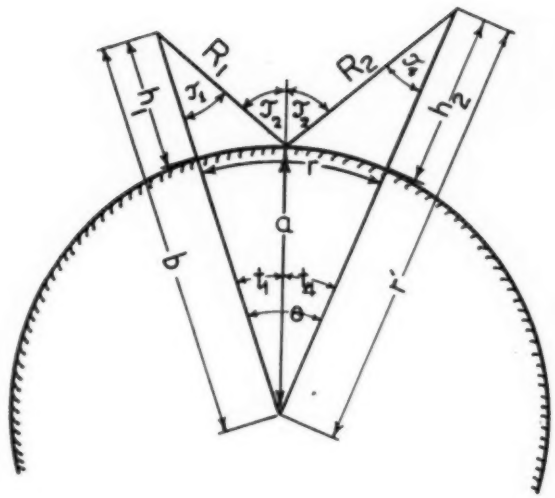


FIG. 9. The ray reflected from the surface of the earth.

We shall now compare (a) with the expression for the reflected wave derived by van d. Pol and Bremmer,[15] and which includes the so-called divergence factor:

$$\Psi_1'=\frac{\exp[i\omega t-ip_3(k^*)]R(k^*)a(\sin\tau_2\cos\tau_2)^{\frac{1}{2}}}{[br'\sin\theta(R_1r'\cos\tau_4+R_2b\cos\tau_1)]^{\frac{1}{2}}}, \qquad (c)$$

where the quantities are defined in Fig. 9. We have, to within terms of the order q^2h^2,

$$1+qh_1=b^2/a^2;\ \ 1+qh_2=r^2/a^2,\ \ q=2/a,$$
$$(1+qh_1-x^2)^{\frac{1}{2}}=(b/a)\cos\tau_1,\ \ (1+qh_2-x^2)^{\frac{1}{2}}=(r'/a)\cos\tau_4,$$
$$x=\sin\tau_2,\ R_1=b\cos\tau_1-a\cos\tau_2,\ R_2=r'\cos\tau_4-a\cos\tau_2, \qquad (d)$$

$$P=2(R_1\cos\tau_4\cos t_1+R_2\cos\tau_1\cos t_4)/(a\cos\tau_1\cos\tau_2\cos\tau_4). \qquad (e)$$

Substituting in (a) we get

$$\Psi_1=\frac{\exp[i\omega t-ip_3(k^*)]a(\sin\tau_2\cos\tau_2)^{\frac{1}{2}}}{[br'r(R_1\cos\tau_4\cos t_1+R_2\cos\tau_1\cos t_4)]^{\frac{1}{2}}}. \qquad (f)$$

Since $\sin\theta\simeq r/a$, and $\cos t_1$ and $\cos t_4$ are close to unity, expressions (c) and (f) agree within the region of validity of the earth-flattening approximation.

[15] Reference 13, p. 844, and reference 2, p. 266.

Letters to the Editor

Linearized Supersonic Flow through Ducts

H. F. Ludloff and F. Reiche
College of Engineering, New York University, New York, New York
April 22, 1947

In various aerodynamical problems, it is desirable to be able to compute the supersonic field of flow through ducted bodies by an analytic, approximate method, which yields quick insight into essential flow features as related to the shape of the nozzle wall. The linearized theory of supersonic flow suggests solutions that can be produced by spreading "supersonic sources" of various intensities over the surface of a hollow skeleton cylinder. If the source density is assumed constant over this surface, the wall contour of a converging diverging nozzle results. With a source intensity which increases linearly along the generatrices of the cylinder, one obtains ducts of gradually decreasing cross section. In the first case, one obtains a duct with a dull rim, and at the entrance a compression shock which has the shape of a double Mach cone, subtended over the rim as the base. This shock is weak near the rim, but becomes stronger toward the vertex and remains particularly strong over the whole surface of the second Mach cone. The linear source distribution produces a knife-sharp rim, as a consequence of which the shocks can be shown to disappear. A certain combination of these two source distributions hinders the formation of shocks, and produces a converging diverging nozzle, whose inside diameter increases at the exit section to the same magnitude as it had at the entrance section. In this way no wave drag results.

Derivation of these results is based on computing the velocity potential from which the velocity and stream function can be derived:

$$\varphi=-2d\cdot\int_0^{\omega^*}d\omega\int_0^{x-\beta R}d\xi\cdot\frac{f(\xi)}{[(x-\xi)^2-\beta^2R^2]^{\frac{1}{2}}}.$$

Here $f(\xi)$ is intensity of source layer at point with coordinates ξ (along generatrix) and ω (azimuth) on surface of cylinder; $x-\xi$ and R, axial and radial components of pseudo-distance between source point and point under consideration. $R^2=d^2+r^2-2d\cdot r\cdot\cos\omega$; d radius of cylinder, r radial distance between axis and point under consideration.

$$\omega^*=\cos^{-1}\frac{d^2+r^2-x^2/\beta^2}{2d\cdot r};\quad \beta^2=M^2-1.$$

(I). For constant source density ($f(\xi)=A$), we get:

$$\varphi=-2d\cdot A\cdot\int_0^{\omega^*}d\omega\cdot\log\frac{\beta R/x}{1-[1-(\beta R/x)^2]^{\frac{1}{2}}}. \tag{1}$$

Differentiating and introducing $R^2=\rho$ instead ω as integration variable:

$$v_r=\frac{\partial\varphi}{\partial r}=\frac{xd\cdot A}{\beta r}\int_b^c\frac{d\rho}{\rho}\cdot\frac{(\rho-(a\cdot b)^{\frac{1}{2}})}{[(c-\rho)(a-\rho)(\rho-b)]^{\frac{1}{2}}}, \tag{2}$$

where

$$a=(d+r)^2,\quad b=(d-r)^2,\quad c=x^2/\beta^2.$$

The expression for v_r can be discussed.

(a) at $x/\beta=d-r$, i.e. at front cone, by expansion:

$$v_r=-\pi\cdot A\cdot\left(\frac{d}{r}\right)^{\frac{1}{2}}\left\{1-\frac{3d+r}{4d}\left[\frac{x/\beta-(d-r)}{2r}\right]\cdots\right\}, \tag{2a}$$

which shows discontinuity, $-\pi\cdot A\cdot(d/r)^{\frac{1}{2}}$, of v_r at front cone.

(b) at $x/\beta=d+r$, i.e. at second Mach cone, by forming:

$$\lim_{c\to a}v_r\sim\lim_{c\to a}\frac{xd\cdot A}{\beta r}\int_b^c\frac{d\rho}{\rho}\times\frac{\rho-(a\cdot b)^{\frac{1}{2}}}{[(c-\rho)(a-\rho)(\rho-b)]^{\frac{1}{2}}}\to\log[(x/\beta)^2-(d+r)^2], \tag{2b}$$

which shows coalescence of singularities and logarithmic infinity of v_r, at second Mach cone.

(c) in the far rear region, by expanding after $\beta R/x$:

$$v_r=\frac{\pi\beta^2r\cdot d\cdot A}{x^2}\left\{1+\frac{3}{4}\frac{\beta^2}{x^2}(r^2+2d^2)\cdots\right\}, \tag{2c}$$

which shows that disturbance velocity vanishes asymptotically.

The stream function can be obtained by means of Stokes three-dimensional representation:

$$\psi=-\rho_0\int_{(d-r)\cdot\beta}^{x}r\cdot\frac{\partial\varphi}{\partial r}\cdot dx'. \tag{3}$$

Upon inserting (2) into (3), we can discuss ψ, upon integrating (2a), (2b), and (2c) with respect to x. In this way one can see that ψ

(a) has no discontinuity at front cone, since it is proportional to $[x/\beta-(d-r)]$,

(b) is regular and finite at second Mach cone, since it is an integral over a logarithmic infinity,

(c) vanishes asymptotically as $1/x$.

The streamtubes are computed from the relation $U\cdot\rho_0\cdot(r^2/2)+\psi(x,r)=U\cdot\rho_0\cdot(d^2/2)$; they show a constriction between first and second Mach cone near the entrance, and at infinity an inside diameter of the same magnitude as at the entrance.

(II). For linearly increasing source distribution ($f(\xi)=\bar{A}\cdot\xi$), we get:

$$\varphi=2\bar{A}\cdot d\int_0^{\omega^*}d\omega\left\{-x\cdot\cosh^{-1}\frac{x}{\beta\cdot R}+(x^2-\beta^2\cdot R^2)^{\frac{1}{2}}\right\} \tag{1}$$

and

$$\frac{\partial\varphi}{\partial r}=v_r=\frac{\bar{A}\cdot\beta\cdot d}{r}\int_b^c\frac{d\rho}{\rho}\frac{(c-\rho)^{\frac{1}{2}}(\rho-(a\cdot b)^{\frac{1}{2}})}{[(a-\rho)(\rho-b)]^{\frac{1}{2}}}. \tag{2}$$

Therefore, (a) at $x/\beta=d-r$, i.e. at front cone:

$$v_r\sim\frac{x/\beta-(d-r)}{(d\cdot r)^{\frac{1}{2}}\pi}\to0, \tag{2a}$$

which means that there is no discontinuity of v_r at entrance, i.e. knife-edged rim.

(b) at $x/\beta=d+r$, i.e. at second Mach cone: v_r remains finite; in (II), (2) $(c-\rho)^{\frac{1}{2}}$ appears now in numerator; therefore for $c\to a$ $(c-\rho)^{\frac{1}{2}}$ and $(a-\rho)^{\frac{1}{2}}$ cancel, rather than forming a cumulative singularity.

(c) at far rear, expansion after $\beta R/x$ yields:

$$v_r=-\frac{\bar{A}\pi\cdot\beta^2r\cdot d}{x}\left\{1+\frac{1}{4}\frac{\beta^2}{x^2}(r^2+2d^2)+\cdots\right\}. \tag{2c}$$

As before, the stream function is obtained by integration with regard to x, so that at far rear:

$$\psi=\rho_0\bar{A}\pi\beta^2r^2\cdot d\cdot[\log x+\cdots]. \tag{3c}$$

Asymptotically, ψ becomes logarithmic infinite; the cross section of the duct decreases as $(r/d)^2 \sim 1/(1+c\cdot\log x)$. For practical purposes, the flow remains parallel within a wide range of x, but a considerable wave drag results.

The advantages of both source distributions (absence of shocks; wide exit section) can be combined, by assuming linearly increasing source intensity $\bar{A}\cdot\xi$ from the rim to a distance $\xi = B$, from where a constant source intensity $\bar{A}\cdot B$ continues. Such a source combination can be resolved into a linear source distribution, starting at the rim, and a linear sink distribution, starting a distance B behind. It is easy to see that the source layer produces a conical compression shock, while the sink layer starts a subsequent expansion wave of equal strength, so that no strong shock will develop anywhere and no wave drag result.

Rigorous Solution of Linearized Supersonic Flow Through Ducts

C. E. MACK, JR., *Grumman Aircraft Engineering Corporation*
AND
H. F. LUDLOFF AND F. REICHE, *College of Engineering, New York University, New York*
April 22, 1947

IN the case of constant source strength along a skeleton cylinder, a rigorous solution of the linearized problem of flow through ducted bodies may be obtained in closed form. The method of approach is such that complete elliptic integrals of the first, second, and third kind arise; however, complete integrals of the third kind may be expressed in terms of elliptic integrals of the first and second kinds so that existing tables suffice for numerical computation.

As in the case of the series expansions, the solution assumes different forms depending upon the location of the test point $P(x, r)$. If P is so located that its fore cone cuts some but not all of the generatrices of the skeleton cylinder, then P is said to lie in the front district. If the fore cone of P cuts all the generatrices of the skeleton cylinder, then P lies in the rear district. Each region will be considered separately.

FRONT DISTRICT $\beta(d-r) < x < \beta(d+r)$

If the value of v_r as given by Eq. (I, 2) above is substituted into Eq. (I, 3) for the disturbed stream function, the latter is given by

$$\psi = -\frac{\rho_0 d\cdot A\beta}{2}\int_{c'=b}^{c'=c} dc' \int_{\rho=b}^{\rho=c'} \frac{d\rho}{\rho}\,\frac{(\rho-(ab)^{\frac{1}{2}})}{[(c'-\rho)(a-\rho)(\rho-b)]^{\frac{1}{2}}} \quad (1)$$

where the symbols have the same meaning as previously. This integral may best be handled by reversing the order of integration, *viz:*

$$\psi = -\frac{\rho_0 d\cdot A\beta}{2}\int_{\rho=b}^{\rho=c} \frac{d\rho(\rho-(ab)^{\frac{1}{2}})}{\rho((a-\rho)(\rho-b))^{\frac{1}{2}}}\int_{c'=\rho}^{c'=c}\frac{dc'}{(c'-\rho)^{\frac{1}{2}}}. \quad (2)$$

Carrying out the integration with respect to dc',

$$\psi = -\rho_0 d\cdot A\beta\left[\int_{\rho=b}^{\rho=c}\frac{(c-\rho)^{\frac{1}{2}}d\rho}{((a-\rho)(\rho-b))^{\frac{1}{2}}} - (ab)^{\frac{1}{2}}\int_{\rho=b}^{\rho=c}\frac{(c-\rho)^{\frac{1}{2}}d\rho}{\rho[(a-\rho)(\rho-b)]^{\frac{1}{2}}}\right]. \quad (3)$$

If now the substitution $\rho = b+(c-b)\sin^2\theta$ is made, then after a short reduction the value of ψ is easily seen to be

$$\psi = -2\frac{\rho_0 d\cdot A\beta}{(a-b)^{\frac{1}{2}}}[(a-b)E_1(k) - (a-c-(ab)^{\frac{1}{2}})F_1(k) - c(a/b)^{\frac{1}{2}}\Pi_1(n,k)], \quad (4)$$

where F_1, E_1, and Π_1 are complete elliptic integrals of the first, second, and third kinds, the parameters k and n being given by

$$k^2 = (c-b)/(a-b), \quad 0\le k<1; \quad n=(c-b)/b, \quad 0\le n<\infty.$$

It has been shown by Cayley that the complete elliptic integral of the third kind may be written, in our notation, as

$$\left[\frac{(n+1)(n+k^2)}{n}\right]^{\frac{1}{2}}\Pi_1(n,k) = \frac{\pi}{2}+\left[\frac{n+k^2}{n(n+1)}\right]^{\frac{1}{2}}F_1(k) + F_1(k)F(k',\theta) - F_1(k)E(k',\theta) - E_1(k)F(k',\theta), \quad (5)$$

where $F(k', \theta)$, $E(k', \theta)$ are incomplete elliptic integrals the parameters of which are

$$k'^2 = 1-k^2, \quad \cot^2\theta = n.$$

When the value of Π_1 from Eq. (5) is substituted into Eq. (4) and the dimensionless ratio $Y = (r/d)(0\le Y<1)$ introduced, it follows at once that

$$\psi = -\rho_0 d^2\cdot A\beta\{4Y^{\frac{1}{2}}[E_1(k)-k'^2F_1(k)] - [(1-Y)^2+4k^2Y]^{\frac{1}{2}}[\pi+2\{F_1(k)-E_1(k)\}F(k',\theta) - 2F_1(k)E(k',\theta)]\}. \quad (6)$$

The stream lines may be obtained from this by evaluating the right-hand side for a suitable range of parameters, combining the results with the stream function for the undisturbed stream and interpolating to find the curves along which the total ψ is constant.

In a similar way to that used above and by the same substitution $\rho = b+(c-b)\sin^2\theta$ it is easy to find the velocity components $v_x = \partial\varphi/\partial x$ and $v_r = \partial\varphi/\partial r$. The results are

$$v_x = -\frac{4Ad}{\beta(a-b)^{\frac{1}{2}}}F_1(k), \quad (7)$$

$$v_r = \frac{2d\cdot A(c)^{\frac{1}{2}}}{r(a-b)^{\frac{1}{2}}}\left[F_1(k)-\left(\frac{a}{b}\right)^{\frac{1}{2}}\Pi_1(n,k)\right]. \quad (8)$$

The elliptic integral of the third kind may be removed as before.

REAR DISTRICT $\beta(d+r) < x < \infty$

The determination of the disturbed stream function for the rear region is very similar to that of the front region, so only the final result need be stated. The value of ψ is given by

$$\psi = -\rho_0 d^2\cdot A\beta\left\{\frac{4}{k}Y^{\frac{1}{2}}\left[E_1(k)-\frac{1-Y}{1+Y}k'^2F_1(k)\right] - \left[(1-Y)^2+\frac{4Y}{k^2}\right]^{\frac{1}{2}}[\pi+2\{F_1(k)-E_1(k)\}F(k',\theta) - 2F_1(k)E(k',\theta)] - 4Y^{\frac{1}{2}}+2(1+Y)\tan^{-1}\frac{2Y^{\frac{1}{2}}}{1-Y}\right\}, \quad (9)$$

where now the parameters k, k', θ are different than for the front region. We have here

$$k^2 = (a-b)/(c-b), \quad 0\le k<1; \quad \cot^2\theta = (a-b)/b$$

and of course $k'^2 = 1 - k^2$. The velocity components are essentially the same as before and need not be reproduced.

Although the solution above is for the case of constant source strength, it may be shown that for polynomial source distributions, the stream lines may be obtained by repeated integration of the stream lines for the constant source. For example, the stream lines for a linear source distribution may be obtained by a single graphical or numerical integration from the constant source stream lines.

The Mechanism of Cutting Fluid Action

MILTON C. SHAW
Mechanical Engineering Department, Massachusetts Institute of Technology, Cambridge, Massachusetts
April 23, 1947

IN a recent letter to the editor,[1] Mr. G. P. Brewington questioned the observation of Ernst and Merchant[2] concerning the lack of correlation between the efficiency of a metal cutting fluid and its physical properties. In support of his contention that some relationship does exist between the cutting ratio and the physical properties of the fluid, Mr. Brewington cites the fact that materials having certain end groups present in their chemical structure (such as $-SH$ or $-Cl$) give a high cutting ratio. It is not clear how this observation substantiates Mr. Brewington's point of view. While it is true that the nature of the end group present on a molecule has an influence on its tendency to become adsorbed on a metal surface, the possibility of the cutting fluid reacting chemically with the metal cut should not be overlooked.

Considerable experimental data have been collected in recent years which indicate that certain effective metal cutting fluids do act by reacting chemically with the metal during the cutting operation. These observations support a rather simple theory of cutting fluid action in which the fluid reacts chemically in the neighborhood of the tool point to form a low shear strength solid lubricant which, in turn, prevents metal-to-metal contact between chip and tool. This mechanism of cutting fluid operation has been described[3] as being a chemico-physical action inasmuch as the fluid first reacts to form chemical products which, in turn, prevent metal-to-metal contact by a physical separation of the surfaces.

The afore-mentioned correlation between chemical structure and cutting performance is, from the writer's point of view, evidence of the correctness of the chemico-physical mechanism of cutting fluid action. Direct chemical analysis has established the presence of solid chemical products of low shear strength on the surface of metal cut in the presence of effective cutting fluids. The fact that certain good cutting fluids are also effective when applied to the tool in their vapor form, the tool and work piece being maintained at a temperature in excess of the boiling point of the fluid, is further evidence of the importance of the chemical rather than physical nature of the fluid. Other observations such as the dependence of the effectiveness of a cutting fluid upon the nature of the metal cut and the fact that benzene is a poorer cutting fluid than air can be explained in terms of the chemical aspects of the problem.

Probably the strongest argument in favor of the chemico-physical theory of cutting fluid action is offered by the discovery that metal cutting processes may be used in the manufacture of chemical products. In this case the product is a chemical rather than a finished machine part. Certain conditions which are found to exist at the point of a cutting tool appear to have an important bearing upon the initiation of a chemical reaction. Three such conditions are:

1. high local pressure—a pressure which will approach the hardness of the metal cut;
2. high local temperature—a temperature which will locally approach the melting point of the metal cut;
3. a nascent highly reactive and highly stressed clean metal surface.

These three conditions may be used to advantage in many organo-metallic reactions particularly those reactions which are difficult to start. In such a case the work piece becomes the metal reactant while the cutting fluid takes the form of the liquid reactant. This method of carrying out a chemical reaction has been termed "mechanical activation."[4] While mechanical activation represents an example of cutting fluid action in which the chemical products formed at the interface between the chip and the tool are present on a macroscopic scale, and, indeed, in some instances represent the entire quantity of metal cut, it should be evident that this technique is in reality a limiting case of cutting fluid action.

The writer concurs in Mr. Brewington's belief that the normal primary monohydric alcohols derive their limited effectiveness as cutting fluids for aluminum from chemical action, and has previously published[5] some substantiating data.

The physical properties of a cutting fluid which control the rate and quantity of fluid which can find its way to the cutting point might well influence the over-all performance of a metal cutting fluid. There thus appear to be two distinct phases of the general problem. First, the physical properties of a useful fluid must be such as to insure sufficient mobility that a satisfactory quantity of fluid reach the cutting point. Second, after the fluid has reached the tool point it is necessary that its chemical nature be such as to enable it to react chemically with the metal cut to form a low shear strength solid lubricant. The required mobility of a cutting fluid will depend largely upon the cutting speed, becoming of negligible significance at very low speed. The afore-mentioned tests of Ernst and Merchant were carried out at a very low cutting speed (5.5 inches per minute) and hence no correlation between cutting efficiency and the physical properties of the cutting fluid should be expected. However, if high speed cutting data were to be examined with regard to the physical properties influencing the mobility of the fluid (viscosity and surface tension), some correlation might be found.

[1] G. P. Brewington, "On the Action of Cutting Oils," J. App. Phys. **18**, 260 (1947).
[2] H. Ernst, and M. E. Merchant, "Chip Formation, Friction and High Quality Machined Surfaces," paper in "Symposium on Surface Treatment of Metals," Am. Soc. Metals, Cleveland (1940), pp. 299–335.
[3] M. C. Shaw, "The Chemico-Physical Role of the Cutting Fluid," Met. Prog. **42**, 85 (1942).
[4] M. C. Shaw, "Mechanical Activation," USP 2,416,717 (1947).
[5] M. C. Shaw, "Action of *n*-Primary Alcohols as Metal Cutting Fluids—Alternating Properties with Chain Length," J. Am. Chem. Soc. **66**, 2057 (1944).

New Books

A Chronological History of Electrical Development

NATIONAL ELECTRICAL MANUFACTURERS ASSOCIATION, 155 East 44 Street, New York, 1946. 106 pp. Price $2.

This is an historical record of electricity and the devices and equipment created by the electrical manufacturing industry to utilize it; of a long line of men of science, their initiative, theories, research, experiments, and discoveries, and their contributions to the present age of electricity, electronics, and atomic energy. The book is designed as a reference work for editors, writers, commentators, libraries, and schools, as well as for industry.

While the book touches on such modern electrical applications as television, fluorescent and ultraviolet light, the giant turbo-generator, and the electron microanalyzer for atomic research, its accent is on the early-day men of science and industry and the steps they took to make electrical living possible. Other contributing factors are recorded, among them electrical expositions and congresses, the birth of electrical publications, engineering and trade associations, societies, foundations, and institutes.

Methods of Mathematical Physics

BY HAROLD JEFFREYS AND BERTHA SWIRLES JEFFREYS. Pp. 679+VII. Cambridge, at the University Press; The Macmillan Company, New York, 1947. Price $15.00.

As the authors state in their preface, this book is intended to provide an account of those parts of pure mathematics that are most frequently needed in physics. The manner in which they propose to carry out their treatment is revealed by a statement in the preface that careful analysis is more important in science than in pure mathematics, not less, and that the easiest way to make a statement reasonably plausible is to give a rigorous proof. It is this reviewer's feeling that that intention has been carried out in ample measure and in most adequate style.

Chapter 1 treats the real variable and includes such topics as the Bolzano-Weierstrass (without so naming it) and Heine-Borel theorems, Riemann and Stieltjes integrals, uniform continuity and uniform convergence, and infinite integrals. Chapters 2 and 3 deal with vectors and tensors, respectively, the treatment throughout bearing in mind the student of physics and emphasizing such topics as the strain and stress tensors, Euler's dynamical equations, parallax, and the motion of a top. Chapter 4 is on matrices and contains a treatment of quadratic and Hermitian forms, the Pauli spin matrices, the Eddington and Dirac 4×4 matrices, crystal structure, and electromagnetic theory. In Chapter 5, we find multiple integrals, including line and surface integrals, Green's and Stokes' theorems, flux and circulation. A careful analysis is made of the notions of area and of length. Potential theory is the content of Chapter 6. Poisson's and Laplace's equations, minimal theorems, the Rayleigh-Ritz method of numerical solution are treated in detail. Operational methods and physical applications of the operational method occupy Chapters 7 and 8, respectively. The applications deal with charging of a condenser, the seismograph, resonance, dissipative and gyroscopic systems, and radioactive disintegration. Chapter 9 (numerical methods) includes such topics as Lagrange's interpolation formula, divided differences, the central difference formula, formulas for the value of a definite integral, numerical solution of algebraic and differential equations. A note on graphical methods recommends that they be honored in the breach rather than in observance. They are best avoided entirely, the authors advise us. Chapter 10 treats the calculus of variations and leads up to the Hamilton equations, principle of least action, and Routh's modified Lagrangian function. Chapters 11, 12, and 13 deal with functions of a complex variable and their application to contour integration and conformal representation. The Joukowsky transformations are treated in detail. Fourier series are the subject matter of Chapter 14. Approximation by polynomials, as well as by trigonometric series, is discussed, also the Laplace transform. Chapter 15 is given over to the factorial and related functions. Our authors frown on the gamma-function as a matter of notation and prefer to label it the factorial function. Wallis's formula for the value of π and Stirling's formula for the value of $\log Z!$ are derived. Chapter 16 is devoted to second-order linear differential equations. These are studied as regards ordinary and singular points and are solved in series and by complex integrals. Infinite determinants are touched on in this connection. Asymptotic expansions occupy Chapter 17, with application to such matters as wave velocity, dispersion of water waves, and refraction of a pulse. Potential waves and heat conduction, waves in one dimension and spherical symmetry, conduction of heat in one and three dimensions, and allied physical ideas are treated in Chapters 18, 19, and 20. The many important systems of coordinates are treated in detail. Bessel functions in Chapter 21 and applications of Bessel functions in Chapter 22 are a logical continuation of the subject matter in the three preceding chapters. In Chapter 23 are found the confluent hypergeometric function, its representation in series and by complex integrals, Whittaker's transformation, Hermite and Laguerre polynomials, and Schrödinger's equation for the hydrogen-like atom. Chapter 24 is given over to Legendre and associated functions and their application to potential theory. Chapter 25, on elliptic functions, closes the text.

Thirteen pages of notes follow, numbered according to the sections of the text to which they refer and designed to clarify or amplify such sections. For example, the Heine-Borel theorem, alluded to in Section 1.072 of the text, is proved in the corresponding section of the notes.

The appendix on notation, which completes the book, embodies recommendations on standardizing certain bits of notation—particularly those with regards to ν and φ—which, as is well known, have not been used with uniformity of meaning in the literature.

The legends posted below the chapter headings are often delightful and illuminating. To quote a few: "one by one, or all at once" (multiple integrals), "but all that moveth

doth mutation love" (potential theory), "even Cambridge mathematicians deserve justice" (operational methods), "I have no satisfaction in formulas unless I feel their numerical magnitude" (numerical methods).

Each chapter is followed by a list of problems, in all—if the reviewer's count is correct—230, and culled, in a large measure, from the British collections such as the Mathematical Tripos.

A high degree of thoroughness, such as will satisfy the pure mathematician, with the entire work set in a frame of physical applications, which enliven and give meaning to the mathematics treated, make this book a welcome addition to the literature of mathematical physics.

MAX MORRIS
Case School of Applied Science

Photography by Infrared. Its Principles and Applications

BY WALTER CLARK. Second Edition. Pp. 472+xvii, Figs. 93, 21¼×14 cm. John Wiley and Sons, Inc., New York, 1946.

The London air raids destroyed the printing plates for the first edition of this book, which was published in England in 1939. The author has taken full advantage of this opportunity to revise thoroughly and modernize the presentation. Included are not only many wartime developments both here and abroad, but also more fundamental advances such as the improvement of infra-red sensitizing dyes.

The result is a comprehensive and authoritative account of the subject, and one that will appeal particularly to the amateur photographer, to the scientific research worker, and to the technician interested in possible application of infra-red. The author states that some acquaintance with photographic practice is assumed, although a fair amount of space is devoted to ordinary photographic procedures, with only occasional reference to the modifications required for work in the infra-red.

This material, including a chapter on darkroom practice, one on the characteristics of films, and an appendix on developers, etc., was presumably included to enhance the value of the book as a laboratory guide. However, it also furnishes a good background for discussion of the problems peculiar to the infra-red, and the general reader with little experience in photography will appreciate the complete and logical way in which the subject as a whole is presented.

The introductory chapter is mostly historical, and the inclusion of direct quotations and illustrations from the early publications, both here and elsewhere in the book, is an attractive feature. Incidentally, it is unfortunate that the very first figure, depicting Newton's experiment on the dispersion of white light, is incorrectly drawn. The reviewer was pleasantly surprised to find due credit given to the physicists and astronomers who did the pioneer work in infra-red photography, because all too often this work has been overlooked. The second chapter contains a useful summary of the general practice of infra-red photography, with references to later chapters where the various matters are treated in more detail. This ingenious arrangement, while very suitable for the purpose in hand, leads to a considerable amount of repetition.

In keeping with the objective of discussing principles as well as applications, further chapters deal with the sensitometric characteristics of photographic materials, sensitization for the infra-red, indirect methods on infra-red photography, and sources of infra-red. These subjects are well presented, and embody a wealth of valuable information. The only criticism that might be made is the failure to distinguish clearly at the outset between the effects of the intensity of the light and of the exposure time. Thus "exposure" is used as abscissas for the characteristic curves without stating that it is the product of intensity and time, and the situation is not improved by giving the wrong units in the first place in which they appear. A very good discussion of the failure of the reciprocity law is given later, but the author has followed the almost universal tendency of writers in this field first to give the impression that the two variables are equivalent, and afterwards to correct this impression.

The remainder of the book is devoted to the various applications of infra-red photography, chief among which are differentiation of materials (textiles, printer's inks, paintings, documents), medicine, botany and paleontology, photo-micrography, landscape photography, forest and other surveys, special effects photography, photography in the dark, spectrographic and astronomical photography, and camouflage detection. Penetration of haze and fog is covered both from the theoretical and observational standpoints, and here the author rightly emphasizes that, contrary to the popular impression, infra-red has little advantage where the large particles of natural fogs are concerned. A final chapter deals with the optical properties of materials in the infra-red. It includes data on the transmission and reflection of common materials, on infra-red filters, and an especially valuable discussion of the infra-red focus of lenses.

Rarely does one find a book as well documented as this one. The bibliographies following each chapter frequently contain over one hundred titles, in many languages. Although a few of the latest developments, such as the concentrated arc and the interference filter, are not included, the coverage of the subject is remarkably complete. This, together with the abundance of valuable data and the many fine illustrations, should justify the prediction that the book will be the standard reference work in infra-red photography for some time to come.

J. A. JENKINS
University of California

Introduction to Electron Optics. The Production, Propagation and Focusing of Electron Beams

BY V. E. COSSLETT. Pp. 272, Figs. 155, 16½×25 cm. The Clarendon Press, Oxford, England. 1946.

Introduction to Electron Optics is primarily a textbook based on a series of lectures delivered by the author, at Oxford, as part of a course in electron optics. The book maintains a reasonably good balance between mathematical

manipulation, physical ideas, and experimental results. The treatments of the various topics are not so long or detailed as to result in a ponderous book but still include sufficient material to serve as a good introduction to the general subject.

The subject matter is presented in the following order, with the first five topics taking up about two-thirds of the total space:

1. Electrostatic fields
2. Electrostatic focusing
3. Magnetic focusing
4. Aberrations
5. Production of electron beams
6. Cathode-ray tubes
7. Electron diffraction and electron microscopes
8. Miscellaneous application.

The mathematical treatments of lenses can be readily followed by students with training in mathematics through differential equations and vector analysis. The presentation is clear and no difficulty with the meanings or symbols should be encountered. (A list of mathematical symbols is appended.) The discussions are well illustrated by line drawings and graphs, as well as numerical data.

The book is intended for students of electron optics, although those doing research in this field will likely find it useful. Each chapter contains a list of references for further reading along the lines of the material of the particular chapter. The use of a general bibliography of a few carefully chosen references for each phase of the subject is to be commended.

The section dealing with electron diffraction is very brief as is that concerned with electron microscopy. In these two instances the reference might have been somewhat more extensive, but it is very likely that those who are interested in applied electron optics will have access to other sources. The emphasis is on fundamentals rather than design and use.

In the opinion of the reviewer, this book should serve exceedingly well as a basis for courses in electron optics in American universities.

R. D. HEIDENREICH
Bell Telephone Laboratories, Murray Hill, New Jersey

Relaxation Methods in Theoretical Physics

BY R. V. SOUTHWELL. Pp. 248. Figs. 117, 16×24½ cm. Oxford, at the Clarendon Press, Oxford, England, 1946.

The technique of solving numerically the Laplace equation in one dimension, by reduction to a difference equation and subsequent application of "relaxation" networks, has become a well-known procedure in advanced problems in engineering. Professor R. V. Southwell and his colleagues have contributed notably to this advance in applied mathematics. Their work has been summarized in exceedingly useful form in Professor Southwell's book, *Relaxation Methods in Engineering Science*, published in 1940. It is highly gratifying, therefore, to have at hand a companion volume which treats in a similar way the two-dimensional problems of mathematical physics.

The book under review, *Relaxation Methods in Theoretical Physics*, is devoted to the solution of the partial differential equation $\frac{\partial}{\partial x}\left[\chi\frac{\partial\psi}{\partial x}\right]+\frac{\partial}{\partial y}\left[\chi\frac{\partial\psi}{\partial y}\right]+Z=0$, where χ and Z may be functions of x and y, and $\psi(x, y)$ is the function sought, subject to prescribed boundary conditions. The arrangement of the book is excellent, and presents enough detail so that the interested reader may understand the background theory and carry through the computational techniques to a useful end.

Two chapters are devoted to the development of the approximating difference equations for the case $\chi=$ constant in a form for numerical solution and to the relaxation techniques for square, triangular, and hexagonal networks. Practical hints for the beginner are included. As pointed out by the author, one can gain a clear insight into the method only by doing the examples in detail.

Irregular networks applied to various boundary conditions are considered in Chapter III. Three types of boundary conditions are treated, *viz.*, (1) ψ has a given value on the boundary, (2) $\partial\psi/\partial\nu$, the normal derivative of ψ, has specified values on the boundary, and (3), a combination of these two conditions. In many cases boundaries of complex shape may be reduced to circular or square boundaries by conformal transformation so that computation by simple nets is more readily effected.

Problems of more complexity, in which χ is a function of x and y, are discussed in detail in Chapter V. While problems of this type are more difficult to solve by orthodox methods than those of the plane harmonic type ($\chi=$ constant), they may be solved numerically with only slight modification in the relaxation technique.

The power of the relaxation method becomes fully apparent in its application to problems in which different equations govern the wanted function or in which some part of the boundary conditions must be determined by computation. Such problems are next to insoluble by orthodox methods. The last chapter of the book treats in detail physical problems of the above type.

An appendix contains formulae for numerical integration and differentiation, a bibliography of papers relating to relaxation methods, a list of the problems solved as examples in the text, and large charts illustrating the solutions. These include torsion problems, conformal transformations, oil pressure distributions in bearings, temperature distribution in a piston, percolation of liquids through walls or strata of various types, etc.

In conclusion one may mention that the basic aim of the numerical approach to physical problems, so admirably treated in this book, is to provide a flexible solution commensurate with the accuracy of the experimental data involved. The wide utility of the relaxation method should be made known to all serious students of physics and engineering who are concerned with more than a "handbook" solution to physical problems. In providing these two basic books, Professor Southwell has performed a major service for the advancement of mathematics in engineering.

S. W. MCCUSKEY
Case Institute of Technology

The Mathematical Tables Project

Sponsored by the National Bureau of Standards. Columbia University Press, New York.

In a recently published book, *An Index of Mathematical Tables*, three British authors have erected a monument to those scientists, who from 1596 on, have published the tables reviewed in the index, an estimated one thousand tables. While it is not possible to contemplate this vast array of published tables without experiencing a feeling of gratitude to the many computers whose work is represented, at the same time it is to be observed that the tables reviewed are published in a bewildering assortment of ways. Some are actually unavailable, and many are practically so. Some are to be found buried in periodical literature, and others constitute a few pages in scattered books. Some of the tables appear as isolated volumes, and still others appear in series of volumes.

The fact that the publications of the New York Project appear as a sequence of tables of similar size and binding, and are sold at nominal prices, makes the series of incalculable value to the modern scientist. As the volumes appear, the set becomes more and more the first source to be consulted when a mathematical table is wanted. The writer of this review looks forward to the day when any table of general usefulness will be found among these volumes.

The care that is exercised in the preparation of these tables, in order to prevent the printing of wrong values, can neither be briefly described nor readily appreciated by any but the professional computer.

The writer of this review regrets that the bound volumes of this table have not been imprinted with distinguishing numbers. In the back of each volume is placed a complete list of all of the volumes prepared prior to the date on which that volume was published. In this list each title bears a reference number. These numbers make better references to a particular volume than do the rather long and similar titles. It seems unfortunate that this number does not appear on the outside cover nor even on the title page. A complete list of titles follows in which those titles which have been recently received are starred.

(1) Table of the First Ten Powers of the Integers from 1 to 1000. 80 pp. (1939). Out of print.
(2) Tables of the Exponential Function e^x. XV+535 pp. (1939).
(3) Tables of Sines and Cosines for Radian Arguments. XIX+275 pp. (1940).
(4) Tables of Circular and Hyperbolic Sines and Cosines for Radian Arguments. XVII+405 pp. (1939).
(5) Tables of Sine, Cosine, and Exponential Integrals, Vol. I. XXVII+444 pp. (1940).
(6) Tables of Sine, Cosine, and Exponential Integrals, Vol. II. XXXVII+225 pp. (1940).
(7) Table of Sine and Cosine Integrals for Arguments from 10 to 100. XXXII+185 pp. (1942).
(8) Tables of Probability Functions.
Vol. I. XXVIII+306 pp. (1941) Vol. II. XXI+348 pp. (1941).
(9) Table of Natural Logarithms.
Vol. I. XVIII+505 pp. (1941) Vol. II. XVIII+505 pp. (1941).
Vol. III. XVIII+505 pp. (1941) Vol. IV. XVIII+510 pp. (1941).
(10) Miscellaneous Physical Tables. VII+61 pp. (1941).
(11) Tables of the Moment of Inertia and Section Modulus of Ordinary Angles, Channels, and Bulb Angles, with Certain Plate Combinations. XIII+201 pp. (1941).
(12) Table of Arc tan$x$. XXV+173 pp. (1942).
* (13) Table of Circular and Hyperbolic Tangents and Cotangents for Radian Arguments. XXXVIII+412 pp. (1943).
* (14) Table of the Bessel Functions $J_0(z)$ and $J_1(z)$ for Complex Arguments. XLIV+406 pp. (1943).
* (15) Table of Reciprocals of the Integers from 100,000 through 200,009. VIII+204 pp. (1943).
* (16) Tables of Lagrangian Interpolation Coefficients. XXXVI+394 pp. (1944).
* (17) Table of Arc sin$x$. XIX+124 pp. (1945).
* (18) Tables of Associated Legendre Functions. XLVI+306 pp. (1945).
* (19) Tables of Fractional Powers. XXX+490 pp. (1946).
* (20) Tables of Spherical Bessell Functions. Vol. I. XXVIII+378 pp. (1947).

ORLEY E. BROWN
Case Institute of Technology

Wave Propagation in Periodic Structures. Electric Filters and Crystal Lattices

BY LEON BRILLOUIN. Pp. 247, 5×8¼ in. McGraw-Hill Book Company, Inc., New York, 1946. Price $4.00.

The rapidly increasing rate at which scientific knowledge is accumulating at the present time would appear to forecast an age of greater and greater specialization in the sciences. Such a state of affairs must undoubtedly be viewed with apprehension by many workers in these domains who have derived a great part of the pleasure in their work from the broadness of understanding with which scientific pursuits have supplied them in the past, and who believe that the interplay of ideas from various specialized fields of a subject such as physics is a healthy and stimulating influence on the development of the subject as a whole. That these fears are as yet ungrounded is demonstrated with force and clarity by Professor Brillouin in this small volume of less than 250 pages, wherein a surprisingly small body of mathematical discipline and associated physical ideas are shown to be quite adequate for the description of a wide range of seemingly diverse physical phenomena, running the gamut from mile long electrical transmission lines to the passage of electrons through crystal lattices.

As its title indicates, the book deals with the general aspects of wave propagation through periodic structures, and examples and applications are drawn from many fields in which such phenomena occur: electric and mechanical filters, vibrations of crystal lattices, x-ray diffraction in crystals, and the quantum theory of metals, to quote some leading examples. The characteristic appearance of transmission and attenuation bands for waves propagating in such periodic structures is demonstrated to have its important physical consequences in each of the above phenomena in such a way that one realizes that a comprehension of the phenomena in any one case almost automatically leads to the understanding of the principles involved in the others. The mathematical apparatus required to discuss adequately almost all of the problems which arise is of such simplicity as to require only the barest acquaintance with the calculus as background for understanding the development of the ideas presented.

This volume, unlike many books, is not intended as a reference work but as a key to understanding. No attempt

is made to treat in detail the various phenomena selected as examples for the general ideas presented. These may be found in many other books. But there are few, if any, other books which demonstrate so directly the broader comprehension which can be obtained of these same physical phenomena when one steps back and examines them as a group linked by a common mathematical and physical background. It is in this respect that the true value of the book becomes immediately apparent.

Professor Brillouin is well known through his other books for the clarity of his exposition, and it is indeed fortunate that he has chosen to make available to others his thoughtful understanding of a subject to which he has been an outstanding contributor.

LESLIE L. FOLDY
University of California

New Booklets

North American Philips Company, Inc., 100 East 42 Street, New York 17, New York, has published a 6-page folder (R1063) entitled *Industrial Control with X-Ray Diffraction*. The text describes the principles of x-ray diffraction and how it is performing difficult jobs in modern industrial applications. Typical x-ray diffraction film is shown and explained. Geiger-counter x-ray spectrometer is discussed in connection with qualitative and quantitative analysis work.

North American Philips also announces a new 4-page folder describing the new Norelco air trimmer, which consists of a stator having three concentric cylinders that slide in the spaces between the four concentric rings of the rotor. Free on request.

New Enterprises, Inc., 84 State Street, Boston 9, Massachusetts, has issued a brochure setting forth its objectives and methods of operation. It was incorporated in February, 1946, "for the purpose of making venture capital more readily available for the development of untried business enterprises of a scientific nature. . . . To qualify for consideration, a new venture must be one of pronounced technical characteristics—a new process, a new or improved technique which shows promise of significant future possibilities."

Leeds and Northrup Company, 4902 Stenton Avenue, Philadelphia 44, Pennsylvania, issues quarterly a $10\frac{1}{2} \times 15\frac{1}{2}$-inch bulletin containing numerous photographs and articles describing the use of Leeds and Northrup controls in industry. 8 pages.

Fischer and Porter Company, Department 6F-C, Hatboro, Pennsylvania, has published a new bulletin describing its "Flowrator" instruments for measuring flow rate of liquids and gases. 24 pages. Available on request.

Bakelite Review for April 1947 contains 28 pages of attractive pictures and articles about consumer goods molded of plastic material. Published quarterly by Bakelite Corporation, 30 East 42d Street, New York 17, New York.

The Tube Department of Radio Corporation of America, Harrison, New Jersey, has resumed issuance of *Application Notes*, which were popular before the war. The first postwar Note is called *Input Admittance of Receiving Tubes*. 20 pages.

Another new publication of the Tube Department is Form No. CRPS-102 entitled *RCA Phototubes, Cathode-Ray Tubes, and Special Tubes*. 16 pages, 10 cents per copy.

Public Affairs Committee, Inc., 22 East 38th Street, New York 16, New York, has published a 32-page pamphlet entitled *The Struggle for Atomic Control*. It was written by William T. R. Fox, Associate Professor of Political Science at Yale University. 20 cents per copy.

Corrosion Publishing Company, 1131 Wolfendale Street, Pittsburgh 12, Pennsylvania, is announcing a booklet of 94 pages entitled *Stress Corrosion Cracking of Mild Steel* by James T. Waber and Hugh J. McDonald. A monograph, it is in large part a reprint of a series of articles which appeared in *Corrosion and Material Protection*. It presents a general theory of stress corrosion and the importance of nitrogen was deduced from it. $2 per copy.

Farrand Optical Company, Inc., Bronx Boulevard and East 238th Street, New York 66, New York, has published *Bulletin 801* entitled *Ultra-Violet Monochromators*. Six instruments are pictured and described. 12 pages.

Burrell Technical Supply Company, 1942 Fifth Avenue, Pittsburgh 19, Pennsylvania, offers a 4-page leaflet describing its precision Mixblend laboratory mixer, *Baker's Analyzed* C. P. Potassium Iodide, Taylor slide comparators.

General Aniline and Film Corporation, 270 Park Avenue, New York 17, New York, has issued a new *Carbonyl Iron Powder* booklet. 34 pages. Available on request.

International Resistance Company, 401 North Broad Street, Philadelphia 8, Pennsylvania, announces its *Bulletin C-2*, which gives complete specifications, characteristics, and dimensions for IRC fixed and adjustable power wire wound resistors, ferrule, and bracket types. 6 pages. Available on request.

The General Radio Experimenter for March, 1947, published by General Radio Company, Cambridge, Massachusetts, features an article on search receivers for radar counter-measures. 8 pages. Available on request to engineers, scientists, and technicians writing on company or professional letterhead.

Here and There

Personnel

Charles C. Bidwell, head of the department of physics at Lehigh University, retired July 1 with the rank of Professor Emeritus.

Five new members have been appointed to the technical staff of the Midwest Research Institute, as follows: in inorganic chemistry, Solon D. Fisher; in organic chemistry, Kenneth R. Hoffman; in engineering mechanics, Eugene F. Hamilton, Richard H. Scheil, and Tage Mortensen.

J. Robert Oppenheimer of the University of California has been appointed to succeed Frank Aydelotte in October, 1947, as Director of the Institute for Advanced Study. Dr. Oppenheimer is a graduate of Harvard and received advanced training in Cambridge University and Göttingen, where he took his Ph.D. He has for the last decade been professor of theoretical physics both in the University of California and in the California Institute of Technology. He was director of the laboratory at Los Alamos which perfected the atomic bomb and is a member of various Government committees for the control of atomic energy.

The Institute for Advanced Study has three Schools: Mathematics, which includes mathematical physics; Economics, which includes history and political science; and Humanistic Studies, which consists at the moment largely of Greek archaeology and the history of art. The Institute provides facilities for postdoctorate research. The Doctor's degree is required for admission, and no degrees are given by the Institute itself.

The appointment of Professor Oppenheimer does not indicate any change in the policy of the Institute as regards the subjects which will be pursued. As it happens, Professor Oppenheimer is a scientist who had as an undergraduate a broad humanistic training.

Dr. Aydelotte emphasized very strongly his approval of this appointment. It would have been difficult to find in the United States a man more ideally qualified for the directorship.

The Institute has a strict rule of retirement at sixty-five. Dr. Aydelotte was kept on until sixty-seven by the trustees at the urgent request of the faculty. He will continue to occupy offices at the Institute for his work in connection with the Rhodes Scholarships, the Guggenheim Fellowships, and the American Friends Service Committee.

U. S. Atomic Energy Commission Announces Distribution of "Heavy Water"

Heavy water and deuterium gas are now being made available by the United States Atomic Energy Committee for research purposes within the United States. The abundance of the heavy hydrogen isotope in the material is approximately 99.9 percent. Quantities will be limited to normal research requirements.

The Stuart Oxygen Company of San Francisco, California, will act as contracting agent for distribution. This company began small-scale production and distribution soon after the discovery of deuterium. In the postwar period, because of the increased need for deuterium in research the Stuart Company's production has been far below the demand. The quantities being made available are from stock produced by other operators during the war for the Manhattan District.

The material will be distributed with charges based on the cost of handling and distribution. The cost of production of the material itself will not be included. Distribution in this manner is being effected under authority of the Atomic Energy Act of 1946 which provides for the fostering and assistance of research by the Commission.

Allocation will be handled in a manner similar to that for radio-isotopes. For complete information write to U. S. Atomic Energy Commission, Research Division, Field Operations, Oak Ridge, Tennessee, attention Isotopes Branch.

Special Expedition Studied Eclipse

A B-29 was used by the National Geographic Society-Army Air Forces cooperative expedition on May 20 in connection with the total eclipse of the sun in Brazil. The plane carried a heavy load of scientific instruments for making studies of cosmic rays in the vicinity of the equator.

Other tests performed by the cooperative expedition dealt with Einstein's theory of relativity, under the direction of Professor G. Van Biesbroeck, astronomer of the Yerkes Observatory of the University of Chicago. The principal objective was an effort to determine whether light rays from certain stars are bent as the rays pass near the sun on their way to the earth, as predicted by Einstein in his theory of relativity.

As a result of the influence of the gravitational field of the sun, light is bent, and star images registered on the photographic plates appear to be shifted away from the sun slightly beyond their actual positions. If this shift amounts to 1.75 seconds of arc, it will be additional evidence that Einstein's theory is correct.

As a check on how much the star images are moved on his photographic plates, Professor Van Biesbroeck will photograph the same stars at night six months later from the same spot. At that time the sun will not be present to bend their rays. Measuring the difference, if any, between the positions of the star images on the two photographs taken six months apart will show how much the light rays were bent during the eclipse.

New Observatory at Colorado

The University of Colorado is collaborating with Harvard University on the creation of a new solar observatory devoted to the study of the region of space which lies between the sun and earth. The new station will be an expansion of the present High Altitude Observatory, where Dr. Walter O. Roberts carries on a study of the sun's outer regions with a coronagraph.

According to present plans, the heart of the new set-up will be a large tube, about six feet in diameter and twenty feet long, set in a dome. In the tube will be placed at least

five observing instruments: a large coronagraph, a small coronagraph to provide automatic "tracking" on the sun, another large telescope devoted to observation of short-wave radiation, and two smaller ones to keep continuous moving picture records of routine sun conditions. In addition, other instruments will aid in the study of cosmic rays and paths of meteorites as a means of providing information about the outer reaches of the atmosphere.

Polymer Research Clinics at Brooklyn

Two summer laboratory clinics remain in the series sponsored by the Institute of Polymer Research of the Polytechnic Institute of Brooklyn. From July 21 to 26 a clinic will be held on "Advanced X-Ray Diffraction". Attendance will be limited to about ten, and the fee is $100. Professor I. Fankuchen is in charge. The September 8–12 clinic is on the subject of "Polymerization Techniques". Attendance will be limited to twelve; fee is $150. Professor Herman Mark is in charge.

Microwave Fellowships at Stanford

The Microwave Laboratory, Division of the Physics Department of Stanford University, has announced a number of fellowships, research assistantships, research associateships, and postdoctorate research fellowships for the academic year 1947–1948. Students who are working in the Microwave Laboratory may receive their degrees in either the Physics or Electrical Engineering Departments. Stipends for the four classes of awards range from $1000 to $4000. For further information and application address Professor William W. Hansen, Director, Microwave Laboratory, Stanford University, California.

Matthew Luckiesh Awarded I.E.S. Medal

Matthew Luckiesh, director of the Lighting Research Laboratory, General Electric Company, has been awarded the I.E.S. Gold Medal, given annually by the Illuminating Engineering Society "for meritorious achievement which has conspicuously furthered the profession, art, and knowledge of illuminating engineering." The medal will be presented to Dr. Luckiesh at the Illuminating Engineering Society National Technical Conference to be held September 15–19 in New Orleans.

Symposium on Applied Mathematics

A Symposium on Applied Mathematics will be held at Brown University, Providence, Rhode Island, August 2–4, 1947, under the sponsorship of the American Mathematical Society and the cosponsorship of the American Institute of Physics. The subject of the Symposium will be "Non-linear Problems in the Mechanics of Continua." All persons interested in receiving further information concerning the Symposium should write to Professor W. Prager, Brown University, Providence 12, Rhode Island. Copies of the program and reservation cards will be sent to them three weeks before the Symposium.

Calendar of Meetings

July

11–12	American Physical Society, Stanford University, California
16–19	American Society of Civil Engineers, Duluth, Minnesota

August

2–4	American Mathematical Society (Symposium on Applied Mathematics), Brown University, Providence, Rhode Island
21–23	Society of Automotive Engineers, Los Angeles, California (Transportation and Maintenance Meeting)
26–29	American Institute of Electrical Engineers, San Diego, California

September

1–2	Mathematical Association of America, New Haven, Connecticut
1–4	American Society of Mechanical Engineers, Salt Lake City, Utah
8–12	Instrument Society of America, Chicago, Illinois
15–19	American Chemical Society, New York, New York
15–19	Illuminating Engineering Society, New Orleans, Louisiana
17–18	Society of Automotive Engineers, Milwaukee, Wisconsin
24–27	Conference on Electrical Insulation, Cambridge, Massachusetts

October

2–4	Society of Automotive Engineers, Los Angeles, California
15–18	American Society of Civil Engineers, New Orleans, Louisiana
20–24	American Society for Metals, Chicago, Illinois
21–25	American Chemical Society, California Section, San Francisco, California
23–25	Optical Society of America, Cincinnati, Ohio

November

3–7	American Institute of Electrical Engineers, Chicago, Illinois
6–7	Society of Automotive Engineers, Tulsa, Oklahoma (Fuels and Lubricants Meeting)
28–29	American Physical Society, Houston, Texas

December

1–3	Society of Automotive Engineers, Kansas City, Missouri (Air Transport Engineering Meeting)
1–5	American Society of Mechanical Engineers, New York, New York
26–31	American Association for the Advancement of Science, Chicago, Illinois
29–31	American Physical Society, Chicago, Illinois